

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 09:49:29 ; Search time 4371 Seconds

(without alignments)  
10239.106 Million cell updates/sec

Title: US-09-940-227-15

Perfect score: 1094  
Sequence: 1 gtccacaggaggactgtacc.....tggctttttaaaaaaaaaa 1094

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hg\_hum:\*  
31: em\_hg\_inv:\*  
32: em\_hg\_other:\*  
33: em\_hg\_mus:\*  
34: em\_hg\_pln:\*  
35: em\_hg\_rod:\*  
36: em\_hg\_mam:\*  
37: em\_hg\_vrt:\*  
38: em\_sy:\*  
39: em\_hg\_hum:\*  
40: em\_hg\_mus:\*  
41: em\_hg\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1094	100.0	1094	6	AX472966
2	799	73.0	4520	9	AF151978
3	563	51.5	3007	10	AF320226
4	562.2	51.4	2337	10	AF161714
5	505.8	46.2	1923	10	AB033285
6	392	35.8	99682	9	HS452H17
7	334.4	30.6	178704	9	BX119904
8	305.8	28.0	151973	2	AC133215
9	305.8	28.0	174655	9	AC023347
10	246	22.5	158527	2	AC027640
11	246	22.5	178704	9	BX119904
12	171.4	15.7	141025	10	AB07250
13	171.2	15.6	195026	2	AC112474
14	171.2	15.6	216341	2	AC134727
15	171.2	15.6	323793	2	AC108660
16	164	15.0	2544	3	AK112740
17	154.2	14.1	6982	10	AF411042
18	154.2	14.1	7081	10	AY147186
19	149.4	13.7	2394	6	AR217626
20	149.4	13.7	2394	6	AR217627
21	149.4	13.7	2450	9	AF142501
22	147.8	13.5	2394	6	AR217628
23	147.8	13.5	2394	9	AF117999
24	147.8	13.5	2394	9	AF352733
25	147.8	13.5	2729	9	AF085412
26	147.8	13.5	2863	6	AR154269
27	147.8	13.5	2863	6	AR221682
28	147.8	13.5	2863	6	E38403
29	136.6	12.5	2817	6	AR048966
30	136.6	12.5	2817	6	AR080685
31	136.6	12.5	2817	10	PATC1X
32	107.8	9.9	137095	9	AC135452
33	107.8	9.9	158527	2	AC027640
34	101.4	9.3	453	6	AR217611
35	101.4	9.3	453	6	AR217612
36	99.8	9.1	453	6	AR217610
37	99.8	9.1	453	6	AR217613
38	99.8	9.1	453	6	AR217614
39	99.8	9.1	453	6	AR217615
40	99.8	9.1	453	6	AR217616
41	99.8	9.1	453	6	AR217617
42	86.6	7.9	755	6	AR217602
43	86.6	7.9	755	6	AR217603
44	86.6	7.9	755	6	AR217604
45	86.6	7.9	755	6	AR217605

## ALIGNMENTS

RESULT 1  
AX472966  
LOCUS AX472966  
DEFINITION Sequence 15 from Patent WO0218576.  
ACCESSION AX472966  
VERSION AX472966.1 GI:22207753  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Chen, S.Y., MacIain, R.A., Sun, Y. and Recipon, H.  
TITLE Compositions and methods relating to lung specific genes  
JOURNAL Patent: WO 0218576-A 15 07-MAR-2002;

Diadexus, Inc. (US)  
 Location/Qualifiers  
 1. 1094  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 338 a 172 c 236 g 348 t

ORIGIN

Query Match 100.0%; Score 1094; DB 6; Length 1094;  
 Best Local Similarity 100.0%; Pred. No. 2,5e-225;  
 Matches 1094; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GTTCACAGGAGGACTGCTTACCTTACAGTTGTTATCGANGAAAAATCAATAAGCAGACTA 60
DB 1 GTTCACAGGAGGACTGCTTACCTTACAGTTGTTATCGANGAAAAATCAATAAGCAGACTA 60
QY 61 AAAACAGCAATAGGCTGCTGATTTGATTTGATCGAACCAAGTGTGCTGCCAAAAGA 120
DB 61 AAAACAGCAATAGGCTGCTGATTTGATTTGATCGAACCAAGTGTGCTGCCAAAAGA 120
QY 121 AAAACCCCTGGGCTTTTAAAGCAAGAACAGGGCTCAGAGTCTATTTCACATATA 180
DB 121 AAAACCCCTGGGCTTTTAAAGCAAGAACAGGGCTCAGAGTCTATTTCACATATA 180
QY 181 TTTGTAATTTGATTAACATACATTTGATTTGATTAACATTTGCGCAATTAAGTGTG 240
DB 181 TTTGTAATTTGATTAACATACATTTGATTTGATTAACATTTGCGCAATTAAGTGTG 240
QY 241 ACATTAATACATGCTATTTCTCCCTCTGAAGAAACGATTCACACAAATTCAGATT 300
DB 241 ACATTAATACATGCTATTTCTCCCTCTGAAGAAACGATTCACACAAATTCAGATT 300
QY 301 ATTTCOCAGAGTATGAGAAAGAAATAGGGTTCCCAATTTGGGCTGCTGCTGTTT 360
DB 301 ATTTCOCAGAGTATGAGAAAGAAATAGGGTTCCCAATTTGGGCTGCTGCTGTTT 360
QY 361 GTTTCCTCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 GTTTCCTCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 CCACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 CCACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 CTGGATTTATGAGGGAGACAGATTCAATTAGAGATACAGAAATGATGAGCAAGAG 540
DB 481 CTGGATTTATGAGGGAGACAGATTCAATTAGAGATACAGAAATGATGAGCAAGAG 540
QY 541 GTGATATTTCTGGCTATGCTGAGAGAGCTTGTGTTTGAATTAAGCCTATTCCTTT 600
DB 541 GTGATATTTCTGGCTATGCTGAGAGAGCTTGTGTTTGAATTAAGCCTATTCCTTT 600
QY 601 TGCATATTTATCTGCTCATTTGCTGCAATTTATGCACTAATTAAGGGCAATTCATA 660
DB 601 TGCATATTTATCTGCTCATTTGCTGCAATTTATGCACTAATTAAGGGCAATTCATA 660
QY 661 CCCTGACGAGGAGTGTGCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 CCCTGACGAGGAGTGTGCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 AATTATGGCTATCAATAAATAATTCAGGCTAAAGAAACATCTTTCAAGCCTTATTA 780
DB 721 AATTATGGCTATCAATAAATAATTCAGGCTAAAGAAACATCTTTCAAGCCTTATTA 780
QY 781 TTCTCTGACAGACGCTTCTAATCTGAGGCTCATCTCGGAAACATCTGAGGAAAGATA 840
DB 781 TTCTCTGACAGACGCTTCTAATCTGAGGCTCATCTCGGAAACATCTGAGGAAAGATA 840
QY 841 TAAAGCATGCTAGTCTCTAATAAAGAGGCTGCGCATGAATTAATCTACTGTTAGTGG 900
DB 841 TAAAGCATGCTAGTCTCTAATAAAGAGGCTGCGCATGAATTAATCTACTGTTAGTGG 900

```

QY 901 CAGAAACCCGAGATGAGATCTCATGTAATAAATATATATGATATGATATTTT 960  
 DB 901 CAGAAACCCGAGATGAGATCTCATGTAATAAATATATATGATATGATATTTT 960  
 QY 961 TAGAATAGGGGAGACCCCTTATTTATTTGTGTGTTACTGATAGAAATGTACTACTAT 1020  
 DB 961 TAGAATAGGGGAGACCCCTTATTTATTTGTGTGTTACTGATAGAAATGTACTACTAT 1020  
 QY 1021 GTTCATATAGGCTGATTTTTCCTTACAGAGAAATGCAATTAATAATGTGTT 1080  
 DB 1021 GTTCATATAGGCTGATTTTTCCTTACAGAGAAATGCAATTAATAATGTGTT 1080  
 QY 1081 TTTTAAAAA 1094  
 DB 1081 TTTTAAAAA 1094

RESULT 2  
 AF151978  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

FEATURES  
 source

1. 4520  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="X"  
 /map="Xq23-q24"  
 /tissue\_type="mammary gland"  
 1. 4520  
 /gene="ATB0+"  
 89. 2017  
 /note="transports zwitterionic and cationic amino acids in Na+ and Cl- dependent manner; highest expression in lung, trachea, and salivary gland; lower expression in the mammary gland, stomach, and pituitary gland; very low expression in the colon, prostate, uterus, and testis"  
 /product="amino acid transporter B0+"  
 /protein\_id="AAD49223.1"  
 /db\_xref="GI:5732680"  
 /translation="MDKIKCPSEFKREKRVASSENHYGENDENDRGNMESKSD TILSMIGTAVGIGNWRPRPYLTYSNGGAPLIPYAIMLAGLPIFELECSLGPASL GPVSWRIPLPEFGVITWELISFTVITYNITATISLYMPASOSLPKKNSSWS DKNCRSRPIVHCNVSIVNKGIOELIIONKSMVDNFTCLNGSETIPQGLPSEQYH NKVALORSNGMNETGVIVYLAICLLAMLVGALFKIGSSGKVEFTALPFPVYL ILLVKGATLEGASKGISYIGAGSNFTKLEAEVWMDAAQOIFSLSVANGGLVALS SVNKRKNCSGSAIVVCLNCTSVAFAPALFSLIGHNAHISGEVSOVVSGBDLAF TAYPAPALQIBGCPWSTLPEFMILLIGDSQFAIETITTTIDLPKVMKMRVPI TLGCVLVEILGVCVTOAGIYVWLLIDHFCAGWGLIATAILELGIWYKGNRFE DTEEMIGAKRWLFTMLWRACWFTVTPILLIAFTSLVQFHRPNTGALPYDPMKATLG WCMIVFCIIWIPIMAIIRIKIOAKNIFORLISCRPASNWPBYLEOHNGERIKDWDV



QY	564	GAGCTTCGCTGTTGTAATTAACGCCTAACCTTTGGATTGCAATTTATCGGCATCTG	623
Db	1757	GAGCTTCCTGCTTTGTCTATCAGCTCTATTTCTGTTGTCTGCAATTTCTAGTCTGGTCAC	1816
QY	624	TGCAATTTTCATAGACCTAATTTATGCGCAATTCATACCTGACTGGGGAGTTGCTT	683
Db	1817	TGAAATTTTCACAGACCTGATTTATGCTATATTTCCATATTCCTGAGCTGGGGAGTTGCT	1876
QY	684	GCCTGGCTATGATGTTTTCGCTATTTGATTTGATTCGAATTAAGCTATCAATTAATAA	743
Db	1877	GCCTGGCTATGATGATTTTCTGCTATTTTGGATTTGATTTGATTTGATTTGATTTGAT	1936
QY	744	TTGAGGCTAAAGGAACATCTTTTCACGCTTTATTAAGTTGCTGACACCAAGCTTCTA	803
Db	1937	TTTCAGCGAAGGAACATACCTTACAGGCTATCATTAACCTGCTGTACGACCAAGCTT	1996
QY	804	GGGGTCCATACCTGGAACAACATGCTGGGGAAATATTAAGACATGCTAGTTCCTTAA	863
Db	1997	GGGGTCCATACCTGGAACAACATGCTGGGGAAATATTAAGACATGCTAGTTCCTTAA	2056
QY	864	AAGAGGCTGGCCATGAATAATACCTACTGTTTACTGGCAGCAGAAAACCGGAATGAGAT	923
Db	2057	AGGAGACTGACACGAAATATACCTACTGTTTATGCGAAGTCAAAACCGAATGAGATG	2116
QY	924	---TTGAAAAAATATATGATTGATGTATATGATTTTTTTTATAGATAGGGGACCC	980
Db	2117	CCTTTGAAGAAATATTAATGTTATATTAATGATTTTGTAGAGTGAAGGAATTTTAT	2176
QY	981	TTATTTGCTGTAAACGATAGGAAAAATGACATACTACTTTCATGATAGTAGGT---	1036
Db	2177	TTATTTGATGTATATCAATATAGGAAATATGACATGTCATCTCAGTACATGATGTAA	2236
QY	1037	TTTTTTTCCCTTTAAGCAGGAATGCAATATATAAAATGTGGTTTTTTTAA	1085
Db	2237	TACTTTTTCCTTAATCAACGAATGAATTTTAAATTTTGTGATCATCTCA	2285

RESULT 4				
AF161714				
LOCUS	AF161714	2337 bp	mRNA	linear
DEFINITION	Mus musculus amino acid transporter B0+ (ATB0+) mRNA, complete cds.			
ACCESSION	AF161714			
VERSION	AF161714.1	GI:5732879		
KEYWORDS				
SOURCE				
ORGANISM	Mus musculus (house mouse)			
	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 2337)			
AUTHORS	Sloan,J.L. and Mager,S.			
TITLE	Cloning and functional expression of a human Na(+) and			
	Cl(-)-dependent neutral and cationic amino acid transporter B(0+)			
	J. Biol. Chem. 274 (34), 23740-23745 (1999)			
JOURNAL				
MEDLINE	99376996			
PUBMED	10446133			
REFERENCE	2 (bases 1 to 2337)			
AUTHORS	Revell,L., Sloan,J.L. and Mager,S.			
TITLE	Direct Submission			
JOURNAL	Submitted (22-JUN-1999) Cell and Molecular Physiology, University			
	of North Carolina, CB 7545, 460 MSRB, Chapel Hill, NC 27599, USA			
FEATURES				
SOURCE	1..2337			
	Location/Qualifiers			

gene	1. .2337
	/gene="ATB0+"
CDS	249. .2165
	/gene="ATB0+"

```

/note="similar to the Homo sapiens ATB0+ encoded by the
sequence presented in GenBank Accession Number AF151978"
/codon_start=1
/product="amino acid transporter B0+"
/protein_id="AADA9320.1"
/db_xref="GI:5732880"
/translation="MDRLKCPNPFCKQKRYTASSENHFYGENDEQOEGNNSKSD
YLLSMVGAVGLGVMWRRPYLTITNGGALIPALMALAGLPLVLESLGCFASL
GVPYSWRILPLFGVGGITWVPLISVEFAIYNYNIIAYSLYLFASFQSLPWANCSMA
DENCSRPIVITGCNVNSIGAGEBMENLWMTNNITCNGSEVRFGDLPSEQWYDRT
LQSSGDETFQGVVWVYLALICLLAMLLVWLVGALPKRGISSGVSLPLFPYVLLIL
IRNATLEGASRQGISYIYGAOSNFTKLAEAWKRAQAQIYSSVACGQVLAIVSNKY
FNNGCTPDALIVCLTICLTCLTVSFGFAIFSLGNHAAHISSEVSSVAGSDPLFIAP
EALALDPAGEFASLLEFPMLLTGLSDPAGSIIETITTFQDLFRAKMRNRVPTIDM
CLIFELGLCEVQAGIYVWHLIDHPCFAGSIIETITTFQDLFIWIGKRVFIEDIM
MICKAKWIFPMLNMRKACFVLTPLTILNLSWTSVKNPFRPYADIPYDWCAGALGCMKI
IFCTIWTIMAILIKIYQAEQNIIILRIISCRPSANMGPVLEKHGERYRDMALPAKET
DHEPPTSSTKPE"
BASE COUNT      623 a      552 g      725 t
ORIGIN

```

[illegible]



```

QY 981 TTTATTTGTGTAACTGAAATGAAATGATACATATGTCATGATAGGCT----GA 1036
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2233 TTTATTTGTATGTTATTCATTAATGAGGAATGATACATGTCATGTCATGATGTTAG 2292
QY 1037 TTTTTCCTCATTTAAGACGAATCATATTAATAATGCTGTTT 1081
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2293 TACTTTTTCCTACTTAATCAAGAAATGAAATATTAATTAATGTAATCT 2337

RESULT 5
AB033285 1923 bp mRNA linear ROD 08-APR-2000
LOCUS Mus musculus mCATB0+ mRNA for colonic system B0+ amino acid
DEFINITION transporter, complete cds.
ACCESSION AB033285
VERSION AB033285.1 GI:7544119
KEYWORDS colonic system B0+ amino acid transporter.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
          Mus musculus.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Ugawa, S.
JOURNAL Mouse Colonic System B0+ Amino Acid Transporter
REFERENCE 2 (bases 1 to 1923)
AUTHORS Ugawa, S.
JOURNAL Published Only in Database (2000)

FEATURES
Source
1..1923
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="ddy"
/db_xref="taxon:10090"
1..1923
/gene="mCATB0+"
7..1923
/gene="mCATB0+"
/codon_start=1
/product="colonic system B0+ amino acid transporter"
/protein_id="BAA94300.1"
/db_xref="GI:7544120"
/translation="MDRLKCPNFKCKEYKPKKKEVYASSPNHVGENDENQENSKSD
          YLSMVGAVGLGNVRFPLTYTNGGAFILYATMLALAGLPFLBESLQFSL
          GPNVSVRIILPFGVGTGTVLISVPAITYNYIATSLYLFASFQSLVEMWNCSSWA
          DENCSRPIYTCNVSTIGAGEMFNISWNTNNDLNGSEVFRPQLPEBOYMDVT
          LQNSGMDRTGVVWVYALCLLAMLVGAALFKGKSGKVVYFALPGLVYVLLIL
          IRKATLEGASRGISYVIGASQNTKLREAVMDATQIFSLSVAMGGLVYVLLIL
          FNNNCYSDAIIVCLNCLSVFAGFAIFSLGMAHISGEVSVYSGPDLAFIYK
          EALALQAPGPFNILEFFMLTLGDSOFASITITTTTODLEPKMKMRVPTTGC
          CLIFLLGLCTVAGIYVWHLIDHCAGGILIAILBAGITWIGGRRITEDIEM
          MIKAKRMFLWLRACRPVITPPLLSAILVMSLVKTHRPDIADIPYDGVALGWCMI
          TFCITNPIMATIKIVQABENIILQRIISCRPASNMGPYLEKHGRBYRDMAPAKET
          DHEIPTISGSTKE"
BASE COUNT 503 a 361 c 445 g 614 t
ORIGIN
Query Match 46.2%; Score 505.8; DB 10; Length 1923;
Best Local Similarity 85.9%; Pred. No. 1,le-98;
Matches 561; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

```

```

QY 384 TGACTACAGCTGGAATTTACTGGGTGATCTGATTTGACACACTTCTGCTGATGAGGCA 443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1391 TGACTACAGCTGGAATTTACTGGGTGATCTGATTTGATGATGATCACTTCTGCTGATGAGGCA 1450
QY 444 TTTTAAATTCAGCATATCTGAGCTACTTGGAAATCATCTGATTTATGAGGCAACAGAT 503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1451 TTTTAAATTCAGCATATCTGAGCTACTTGGAAATCATCTGATTTATGAGGCAACAGAT 1510
QY 504 TCATTGAGATACAGAAATGATGATTTGAGACCAAGAGGTGATATTCGTGCTATGAGGCA 563
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1511 TCATTGAGATATGAGAAATGATGATTTGAGACCAAGAGGTGATATTCGTGCTATGAGGCA 1570
QY 564 GAGCTTCTGCTGTTGTAATTAACGCTTATCTTTGATGCAATATTTATCTGCTATG 623
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1571 GAGCTTCTGCTGTTGTAATTAACGCTTATCTTTGATGCAATATTTATCTGCTATG 1630
QY 624 TGCATTTTCATAGACCTAATTTATGAGGCAATTCACCTGACCTGGGAGTTCCTTAC 683
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1631 TGCATTTTCATAGACCTGATTTATGCTGATATTCATATCTGACCTGGGAGTTCCTTAC 1690
QY 684 GCTGCTGATGATTTGTTCTGCAATTTTGGATTCGCAATTTATGCTATCAATAAATA 743
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1691 GCTGCTGATGATTTGTTCTGCAATTTTGGATTCGCAATTTATGCTATCAATAAATA 1750
QY 744 TTCAGGCTAAGAAACATCTTTCACGCTTATTAAGTCTGACAGACGCTTCACT 803
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1751 TTCAGGCTAAGAAACATCTTTCACGCTTATTAAGTCTGACAGACGCTTCACT 1810
QY 804 GGGGTCATACCTGGAACACATCTGAGGAAATATTAAGCATGATGATGCTCTAATA 863
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1811 GGGGTCATACCTGGAACACATCTGAGGAAATATTAAGCATGATGATGCTCTAATA 1870
QY 864 AAGAGCTGCGCCATGAATATCTGATGATGAGGCAAGAAACCGGAATGA 916
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1871 AAGAGCTGCGCCATGAATATCTGATGATGAGGCAAGAAACCGGAATGA 1923

RESULT 6
HS452H17 99682 bp DNA linear PRI 05-MAR-2003
LOCUS Human DNA sequence from clone RP3-452H17 on chromosome Xq22.1-23,
          complete sequence.
ACCESSION 296810
VERSION 296810.1 GI:2276315
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 99682)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
JOURNAL Graftman, D.
          Direct Submision
          Submitted (05-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
          Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
          humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
          On Jul 24, 1997 this sequence version replaced gi:2198470.
COMMENT ----- Genome Center
          Center: Wellcome Trust Sanger Institute
          Center code: SC
          Web site: http://www.sanger.ac.uk
          Contact: humquery@sanger.ac.uk
          -----
          During sequence assembly data is compared from overlapping clones.
          Where differences are found these are annotated as variations
          together with a note of the overlapping clone name. Note that the
          variation annotation may not be found in the sequence submission
          corresponding to the overlapping clone, as we submit sequences with
          only a small overlap as described above.
          This sequence was finished as follows unless otherwise noted: all
          regions were either double-stranded or sequenced with an alternate
          chemistry or covered by high quality data (i.e., phred quality >=
          30); an attempt was made to resolve all sequencing problems, such

```

as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Bm.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone configs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChX>

RP3-452H17 is from the library RPci-3 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

#### FEATURES

##### source

1..99682  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/map="q22.1-23"  
/clone="RP3-452H17"  
/clone\_lib="RPci-3"

BASE COUNT 28716 a 17592 c 17583 g 35791 t

##### ORIGIN

##### Query Match

Best Local Similarity 100.0%; Pred. No. 2e-74; Length 99682;  
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTCACAGGGGAGCTGTACCTTACAGTGTGTTATGATGAAAAATCATATAACGACACTA 60  
D 4240 GTTCACAGGGGAGCTGTACCTTACAGTGTGTTATGATGAAAAATCATATAACGACACTA 4299  
QY 61 AAAACAGCAATAGCGGTGATTTGATTCATCAGAACCAAGGCTGATGCGCAAGA 120  
D 4300 AAAACAGCAATAGCGGTGATTTGATTCATCAGAACCAAGGCTGATGCGCAAGA 4359  
QY 121 AAAACCGCTGTGGTGTAAAAAGCAAGAAAGGCGCTCAGCTATATTTTCAACATATA 180  
D 4360 AAAACCGCTGTGGTGTAAAAAGCAAGAAAGGCGCTCAGCTATATTTTCAACATATA 4419  
QY 181 TTGTGAATTTGATTAACATACATGTAATCTGTGATTACAGATTGCAATTAAGTTG 240  
D 4420 TTGTGAATTTGATTAACATACATGTAATCTGTGATTACAGATTGCAATTAAGTTG 4479  
QY 241 ACATATTAACATGATCTCCCTCTGAAAGAAACGATCACAACAATTCAGAATTT 300  
D 4480 ACATATTAACATGATCTCCCTCTGAAAGAAACGATCACAACAATTCAGAATTT 4539  
QY 301 ATTTCCCAAGTGAAGAAATGAGGGTCCCATTAATTTGGGCTGCTGCTTTT 360  
D 4540 ATTTCCCAAGTGAAGAAATGAGGGTCCCATTAATTTGGGCTGCTGCTTTT 4599  
QY 361 GTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 392  
D 4600 GTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4631

#### RESULT 7

##### BX119904/c

LOCUS BX119904 178704 bp DNA linear PRI 10-FEB-2003  
DEFINITION Human DNA sequence from clone RP11-232D9 on chromosome X, complete  
sequence.

ACCESSION BX119904  
VERSION BX119904.4 GI:28193290

##### KEYWORDS

HTG.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

#### REFERENCE

1 (bases 1 to 178704)  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
AUTHORS Howden, P.  
TITLE Direct Submission  
JOURNAL Submitted (06-FEB-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk) Clone requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk) On Feb 1, 2003 this sequence version replaced gi:27902176.

#### COMMENT

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Bm.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone configs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChX>

RP11-232D9 is from the library RPci-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6.

#### FEATURES

##### source

1..178704  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/clone="RP11-232D9"  
/clone\_lib="RPci-11.1"

BASE COUNT 56317 a 32304 c 35259 g 56554 t

##### ORIGIN

##### Query Match

Best Local Similarity 90.6%; Pred. No. 4.6e-62; Length 178704;  
Matches 356; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1 GTTCACAGGGGAGCTGTACCTTACAGTGTGTTATGATGAAAAATCATATAACGACACTA 60  
D 65147 GTTCACAGGGGAGCTGTACCTTACAGTGTGTTATGATGAAAAATCATATAACGACACTA 65088  
QY 61 AAAACAGCAATAGCGGTGATTTGATTCATCAGAACCAAGGCTGATGCGCAAGA 120  
D 65087 AAAACAGCAATAGCGGTGATTTGATTCATCAGAACCAAGGCTGATGCGCAAGA 65028  
QY 121 AAAACCGCTGTGGTGTAAAAAGCAAGAAAGGCGCTCAGCTATATTTTCAACATATA 180  
D 65027 AAAACCGCTGTGGTGTAAAAAGCAAGAAAGGCGCTCAGCTATATTTTCAACATATA 64968  
QY 181 TTGTGAATTTGATTAACATACATGTAATCTGTGATTACAGATTGCAATTAAGTTG 240  
D 64967 TTGTGAATTTGATTAACATACATGTAATCTGTGATTACAGATTGCAATTAAGTTG 64908  
QY 241 ACATATTAACATGATCTCCCTCTGAAAGAAACGATCACAACAATTCAGAATTT 300

```

Db      64907 AATATATACATGCTATTTTCCCTGAGAGATGATCAAAACAGATTCAAGATT 64848
QY      301 ATTTCCCAAGATGAGAAATAGAGTTCCCACTACTGTCGCTGCTGCTTT 360
Db      64847 ACTTCCCAAGATGAGAAATAGATGATGATGCTTGGCTGCTGCTGCTTT 64788
QY      361 GTTCTCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 392
Db      64787 GTTCTCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 64756

RESULT 8
AC133215
LOCUS   AC133215 151973 bp DNA linear HTG 07-SEP-2002
DEFINITION Homo sapiens chromosome UNK clone RP13-792C11, WORKING DRAFT
SEQUENCE AC133215
VERSION  AC133215.1 GI:22758609
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE   Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 (bases 1 to 151973)
  Waterston,R.H.
  The sequence of Homo sapiens clone
  Unpublished
  2 (bases 1 to 151973)
  Waterston,R.H.
  Direct Submission
  Submitted (07-SEP-2002) Genome Sequencing Center, Washington
  University School of Medicine, 4444 Forest Park Parkway, St. Louis,
  MO 63108, USA

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
        Project Information
Center project name: H.FH0792C11
----- Summary Statistics -----
Sequencing vector: M13; 08
Chemistry: Dye-terminator Big Dye; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 149763 bases at least Q40
Consensus quality: 149919 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
-----
1 1119: contig of 1119 bp in length
* 1120: gap of unknown length
* 1220: contig of 1098 bp in length
* 2318: gap of unknown length
* 2418: contig of 1966 bp in length
* 4384: gap of unknown length
* 4484: contig of 2036 bp in length
* 6520: gap of unknown length
* 6620: contig of 3896 bp in length
* 10516: gap of unknown length
* 10616: contig of 7512 bp in length
* 18127: gap of unknown length
* 18128:

```

```

* 18228 29110: contig of 10883 bp in length
* 29111 29210: gap of unknown length
* 29211 47743: contig of 18533 bp in length
* 47744 47843: gap of unknown length
* 47844 69544: contig of 21701 bp in length
* 69545 69645: gap of unknown length
* 69645 92308: contig of 22664 bp in length
* 92309 92408: gap of unknown length
* 92409 123518: contig of 31110 bp in length
* 123519 123619: gap of unknown length
* 123619 151973: contig of 28355 bp in length.

FEATURES
    source
        1. 151973
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="UNK"
            /clone="RP13-792C11"
            1. 1119
                /note="assembly_name:Contig1"
            1220. 2317
                /note="assembly_name:Contig5"
            2418. 4383
                /note="assembly_name:Contig8"
            4484. 6519
                /note="assembly_name:Contig9"
            6620. 10515
                /note="assembly_name:Contig10"
            10616. 18127
                /note="assembly_name:Contig11"
            18228. 29110
                /note="assembly_name:Contig12"
            29211. 47743
                /note="assembly_name:Contig13"
            clone_end:17
                /note="assembly_name:Contig14"
            47844. 69544
                /note="assembly_name:Contig15"
            69645. 92308
                /note="assembly_name:Contig16"
            92409. 123518
                /note="assembly_name:Contig17"
            123619. 151973
                /note="assembly_name:Contig17"
            clone_end:SP6
                /note="assembly_name:Contig17"
            vector_side:left"

BASE COUNT  47782 a 28922 c 29161 g 45007 t 1101 others
ORIGIN
Query Match      28.0%; Score 305.8; DB 2; Length 151973;
Best Local Similarity 76.1%; Pred. No. 6.8e-56;
Matches 451; Conservative 0; Mismatches 117; Indels 25; Gaps 5;

QY      375 TCGTGTGTGATCAGGCTGGAATTTACTGGGTCATGCTGATTCAGCACTTGTGCTG 434
Db      146271 TCACCGGTGTGCTGAGCTGGAATTTACT -GGCTCATGCTGCTGCACTTGTGCTG 146329
QY      435 GATGGGCAATTTAATGACGCTAATCTGAGCTAGTTCGATTCATTCGATTTATGAG 494
Db      146330 AACGGCAATTTGATGCAACTATACCTGGAATATAGGAATCATCTGGGTTATGAG 146389
QY      495 GGAACGATTCATTCAGATCAGAAATGATGATGAGCAAGAGGTGATATTCGCG 554
Db      146390 GGAACGATTCCTTGAAGATATAAATATATATAT ---GGAACAGAGGTGATATTCGCG 146446
QY      555 TATGTGGAGACCTGCTGCTGTTG-----AATTAGCCTATCCCTTTTG 598
Db      146447 TCACTGGAAACTGCTGCTGTTGCTATTAATGATGACATTAATGCTATTTT 146506
QY      599 ATTGCAATTTATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658
Db      146507 ATTGCAATTTTATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146566

```



```

repeat_region 17454. .18620
/rpt_family="L1"
repeat_region 18777. .19121
/rpt_family="MALR"
repeat_region 19259. .19367
/rpt_family="CRL"
repeat_region 20344. .20420
/rpt_family="MIR"
repeat_region 20595. .20900
/rpt_family="Alu"
repeat_region 21011. .21130
/rpt_family="MIR"
repeat_region 22367. .22439
/rpt_family="MALR"
repeat_region 22440. .22728
/rpt_family="Alu"
repeat_region 22729. .23277
/rpt_family="MALR"
repeat_region 23300. .23387
/rpt_family="L2"
repeat_region 23388. .23785
/rpt_family="MALR"
repeat_region 23786. .23882
/rpt_family="L2"
repeat_region 24530. .24733
/rpt_family="L1"
repeat_region 24970. .25027
/rpt_family="L2"
repeat_region 25974. .26806
/rpt_family="L1"
repeat_region 26832. .27067
/rpt_family="L1"
repeat_region 28690. .29052
/rpt_family="L2"
repeat_region 29118. .29663
/rpt_family="L1"
repeat_region 30891. .31010
/rpt_family="L1"
repeat_region 31011. .31308
/rpt_family="Alu"
repeat_region 31309. .31350
/rpt_family="L1"
repeat_region 31351. .31716
/rpt_family="MERL-type"
repeat_region 31717. .31830
/rpt_family="L1"
repeat_region 32007. .32287
/rpt_family="L1"
repeat_region 32341. .32413
/rpt_family="L1"
repeat_region 32453. .32998
/rpt_family="ERV1"
repeat_region 33534. .33712
/rpt_family="MERS3"
repeat_region 33771. .34069
/rpt_family="Alu"
repeat_region 34517. .34604
/rpt_family="L1"
repeat_region 34848. .34980
/rpt_family="L1"
repeat_region 34981. .35528
/rpt_family="L1"
repeat_region 35628. .35872
/rpt_family="L1"
repeat_region 35873. .36922
/rpt_family="L1"
repeat_region 36923. .37302
/rpt_family="L1"
repeat_region 37678. .37925
/rpt_family="MIR"
repeat_region 37939. .38108
/rpt_family="MERL-type"
repeat_region 39114. .39713

```

```

repeat_region /rpt_family="L1"
39707. .39988
/rpt_family="L1"
repeat_region 40613. .40725

Query Match 28.0% Score 305.8; DB 9; Length 174655;
Best Local Similarity 76.1% Pred. No. 6,66-56;
Matches 451; Conservative 0; Mismatches 117; Indels 25; Gaps 5;

QY 375 TCGTCGTGTCACGACGCTGGAATTTACTGAGTTCATGATGACCACTTCGTGCTG 434
D 122834 TCACCGCTGCTGCTGACGCTGGAATTTAGT-GGCTCATCTGACGCGCACTTCGTACTG 122776
QY 435 GATGGGCAATTTTATTTATGACGCTTACTGAGCTAGTGTGATCATGATCATGATTTATGAG 494
D 122775 AACGGACATTTTATTTATGACACTTACTGACACTAATAGATATATCTGGTTTATGAG 122716
QY 495 GGAACAGATTCATTGAGGATACAGAAATGATGATTTGAGCAAAAGGTGATTTATTCGCG 554
D 122715 GGAACAGATTCCTTGAAGATATATAAATGATTAAT---GGAACAGAGGTGATTTCTCGC 122659
QY 555 TATGCTGAGAGCTTCTGCTGTTGT-----AATTACGCTTATCTTTTG 598
D 122658 TGCAGTGAATAACTTCTGCTGTTGTCATTAATCTATATGACATTAATGCTTATTTT 122599
QY 599 ATTGCAATATTTTATCTGCTCATTTGCTGCAATTTCAATTCATGACCTAATTTAGGCGCAATTC 658
D 122598 ATTGCAATTTTATCTGACACTGATGATTAATATCATATGACTTGATTTGCCAANAATCTTG 122539
QY 659 TACCTGACCTGGGAGTGTCTTAAAGCTGTGATGATGATTTCTGCAATTTATTTGGATT 718
D 122538 TGCCTGACCTGAGAAAGTGTCTCAAGCTGTGATGATTTCTGATTA---TTGCATT 122480
QY 719 CCAATTATGCTATCATTAATAATATCAGCTTAAGAAACATCTTCAAGCCTTATA 778
D 122479 TCAATTATGCTATCATTAATAATATCAGCTTAAGAAACATCTTCTGCTGTTGTA 122420
QY 779 AGTTGCTGACAGACGCTTCACTGAGGCTCACTACCTGGAACAACTGCGGAGAAAGA 838
D 122419 AGCTGCTGTAACAGAGTTTCACTGAGGCTCACTACCTGGAACAACTGCGGAGAAAGA 122360
QY 839 TATTAAGACATGTAGTCTTCAAAAAGAGC---CTGGCCATGAATTAACCTACTGTTAG 894
D 122359 TAAAAAGGACAGATGATCTCTTAATTAAGAGAGCTGAGCAGAGAAATTCCTACTGTTG 122300
QY 895 TGGCAGAGAAACCGGAATGAGATCTCATGTAATAAATAATGATGATTA 947
D 122299 TGGCAGAGAAACCGGAATGAGATCTCATGTAATAAATAATGATGATTA 122247

RESULT 10
AC027640 158527 bp DNA linear HTG 01-SEP-2000
LOCUS Homo sapiens chromosome X clone RP11-714C1, WORKING DRAFT SEQUENCE,
DEFINITION 33 unordered pieces.
ACCESSION AC027640
VERSION AC027640.3 GI:9958776
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 158527)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL unpublished
REFERENCE 2 (bases 1 to 158527)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Sep 1, 2000 this sequence version replaced gi:8318566.

```

Genome Center	
-----	
Center:	Washington University Genome Sequencing Center
Center code:	WUGSC
Web site:	<a href="http://genome.wustl.edu/gsc/index.shtml">http://genome.wustl.edu/gsc/index.shtml</a>
Project Information	
Center project name:	H_NH071AC01
-----	
Summary Statistics	
-----	
Sequencing vector:	ML3; 100%
Chemistry:	dye-terminator; 100% of reads
Chemistry:	dye-terminator; Big Dye; 0% of reads
Assembly program:	Phrap; version 0.990319
Consensus quality:	148926 bases at least Q40
Consensus quality:	151403 bases at least Q30
Consensus quality:	152426 bases at least Q20
Insert size:	16300; agarose-fp
Insert size:	15327; sum-of-contigs
Quality coverage:	5.23 in Q20 bases; sum-of-contigs
-----	
* NOTE: This is a 'working draft' sequence. It currently	
* consists of 33 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	
-----	
1	3335: contig of 3335 bp in length
3336	3335: gap of unknown length
3436	5836: contig of 2401 bp in length
5837	5936: gap of unknown length
5937	8911: contig of 2975 bp in length
8912	9011: gap of unknown length
9012	11752: contig of 2741 bp in length
11753	11852: gap of unknown length
11853	14013: contig of 2161 bp in length
14014	14113: gap of unknown length
14114	17278: contig of 3165 bp in length
17279	17378: gap of unknown length
17379	19490: contig of 2112 bp in length
19491	19590: gap of unknown length
19591	21828: contig of 2238 bp in length
21829	21928: gap of unknown length
21929	24934: contig of 3006 bp in length
24935	25034: gap of unknown length
25035	28319: contig of 3285 bp in length
28320	28419: gap of unknown length
28420	31582: contig of 3163 bp in length
31583	31682: gap of unknown length
31683	34658: contig of 2976 bp in length
34659	34758: gap of unknown length
34759	38924: contig of 3566 bp in length
38925	38924: gap of unknown length
38425	41443: contig of 2819 bp in length
41444	41343: gap of unknown length
41344	43546: contig of 2203 bp in length
43547	43646: gap of unknown length
43647	46071: contig of 2425 bp in length
46072	46171: gap of unknown length
46172	63170: contig of 16999 bp in length
63171	63370: gap of unknown length
63371	67563: contig of 4293 bp in length
67564	67563: gap of unknown length
67564	72467: contig of 4804 bp in length
72468	72467: gap of unknown length
72568	72567: contig of 4985 bp in length
77553	77652: gap of unknown length
77653	83764: contig of 6112 bp in length
83764	83864: gap of unknown length
83865	90080: contig of 6216 bp in length
90081	90180: gap of unknown length

	*	90181	97167:	contig of 6987 bp in length
	*	97168	97267:	gap of unknown length
	*	97268	106868:	contig of 9601 bp in length
	*	106968	106968:	gap of unknown length
	*	106969	114304:	contig of 7336 bp in length
	*	114405	114404:	gap of unknown length
	*	114405	129608:	contig of 15204 bp in length
	*	129609	129708:	gap of unknown length
	*	129709	146858:	contig of 17150 bp in length
	*	146859	146958:	gap of unknown length
	*	146959	148941:	contig of 1983 bp in length
	*	148942	149041:	gap of unknown length
	*	149042	150952:	contig of 1911 bp in length
	*	150953	151052:	gap of unknown length
	*	151053	152528:	contig of 1476 bp in length
	*	152529	152628:	gap of unknown length
	*	152629	154332:	contig of 1694 bp in length
	*	154423	154422:	gap of unknown length
	*	154423	156637:	contig of 2215 bp in length
	*	156638	156737:	gap of unknown length
	*	156738	158527:	contig of 1790 bp in length
FEATURES			Location/Qualifiers	
source			1..158527	
			/organism="Homo sapiens"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:9606"	
			/chromosome="X"	
			/clone="RP11-714C1"	
			1..3335	
misc_feature			/note="assembly_name:Contig10"	
misc_feature			3436..5836	
misc_feature			/note="assembly_name:Contig11"	
misc_feature			5937..8911	
misc_feature			/note="assembly_name:Contig12"	
misc_feature			9012..11752	
misc_feature			/note="assembly_name:Contig13"	
misc_feature			11853..14013	
misc_feature			/note="assembly_name:Contig14"	
misc_feature			14114..17278	
misc_feature			/note="assembly_name:Contig15"	
misc_feature			17379..19450	
misc_feature			/note="assembly_name:Contig16"	
misc_feature			19591..21828	
misc_feature			/note="assembly_name:Contig17"	
misc_feature			21929..24934	
misc_feature			/note="assembly_name:Contig18"	
misc_feature			25035..28319	
misc_feature			/note="assembly_name:Contig19"	
misc_feature			28420..31582	
misc_feature			/note="assembly_name:Contig20"	
misc_feature			31683..34658	
misc_feature			/note="assembly_name:Contig21"	
misc_feature			34759..38324	
misc_feature			/note="assembly_name:Contig22"	
misc_feature			38425..41243	
misc_feature			/note="assembly_name:Contig23"	
misc_feature			41344..43546	
misc_feature			/note="assembly_name:Contig24"	
misc_feature			43647..46071	
misc_feature			/note="assembly_name:Contig25"	
misc_feature			clone_end:36	
misc_feature			vector_side:left	
misc_feature			46172..63170	
misc_feature			/note="assembly_name:Contig26"	
misc_feature			clone_end:17	
misc_feature			vector_side:right	
misc_feature			63271..67563	
misc_feature			/note="assembly_name:Contig26"	
misc_feature			67664..72457	
misc_feature			/note="assembly_name:Contig27"	
misc_feature			72568..77552	
misc_feature			/note="assembly_name:Contig28"	
misc_feature			77653..83764	

```

misc_feature /note="assembly_name:Contig29"
83865..90080
misc_feature /note="assembly_name:Contig30"
90181..97167
misc_feature /note="assembly_name:Contig31"
97268..106868
misc_feature /note="assembly_name:Contig32"
106969..114304
misc_feature /note="assembly_name:Contig33"
114405..129608
misc_feature /note="assembly_name:Contig34"
129709..146858
misc_feature /note="assembly_name:Contig35"
146959..148941
misc_feature /note="assembly_name:Contig4"
149042..150952
misc_feature /note="assembly_name:Contig5"
151053..153328
misc_feature /note="assembly_name:Contig6"
152629..154322
misc_feature /note="assembly_name:Contig7"
154423..156637
misc_feature /note="assembly_name:Contig8"
156738..158527
misc_feature /note="assembly_name:Contig9"

BASE COUNT 48002 a 28204 c 31062 g 48053 t 3206 others
ORIGIN
Query Match 22.5%; Score 246; DB 2; Length 158527;
Best Local Similarity 87.0%; Pred. No. 5,1e-43;
Matches 282; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

QY 765 TTCACGCTTTAAAGTTGCTGCAGACAGCTTCTAAGTGGGCTCATACCTGGAACAC 824
DB 107845 TTCAGTGCCTTGAAGTTCGTCAGACAGCTTCTAAGTGGGCTCATACCTGGAACAC 107786
QY 825 ATCTGGGGAAAGATATTAAGCATGAGTTCCTCAAAAAGAGCGCTGGCCATGAATAC 884
DB 107785 GTTCGGGGAGAGATATTAAGCATGATATATCCCAAAAAGAGCATGATGACATAC 107726
QY 885 CTACTGTATAGTGGCAGCAAAACCGAATGAGATCTCATTTGAAAAAAATATATGATGT 944
DB 107735 CTACTGTATAGTGGCAGCAAAACCGAATGAGATCTCATTTGAAAAAAATATATGATGT 107666
QY 945 ATAATGTGAT--TTTTTAAATAGAGGGGACCTTATTTATTTGTGTGTTAACTGATA 1002
DB 107665 ATTAATGTATTTGTTTTTAAAGAAATGGGGAAACTTTTATTTGTTAATGTTAACTGATA 107606
QY 1003 GGAATATGTACATCTATGTTTCATGATAGGTTGTTTCCCATTTAGCGAATGC 1062
DB 107605 GGAATATGTACATCTATTTTCATGATAGTGTGATTTTCTCACATTTAAGCAGGAATGC 107546
QY 1063 AATATAAAAATGTCGTTTTTTAA 1086
DB 107545 AATATAAAAATGTAATTTCTTAA 107522

RESULT 11
LOCUS BX119904 178704 bp DNA linear PRI 10-FEB-2003
DEFINITION Human DNA sequence from clone RP11-232D9 on chromosome X, complete
sequence.
ACCESSION BX119904
VERSION BX119904.4 GI:28193290
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 178704)
AUTHORS Howden,P.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-2003) Wellcome Trust Sanger Institute, Hinxton,

```

```

COMMENT
Cambridgehire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 1, 2003 this sequence version replaced gi:27902176.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
EM, EMBL; SW, SWISSPROT; Tr, TREMBL; WP, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RP11-232D9 is from the library RP11-1.1 constructed by the group
of Pletier de Jong. For further details see
http://www.chori.org/Bacpac/home.htm
VECTOR: pBACE3.6.

FEATURES
source
location/Qualifiers
1..178704
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-232D9"
/cclone.lib="RP11-232D9"

BASE COUNT 56317 a 32304 c 33529 g 36554 t
ORIGIN
Query Match 22.5%; Score 246; DB 9; Length 178704;
Best Local Similarity 87.0%; Pred. No. 5e-43;
Matches 282; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

QY 765 TTCACGCTTTAAAGTTGCTGCAGACAGCTTCTAAGTGGGCTCATACCTGGAACAC 824
DB 83945 TTCAGTGCCTTGAAGTTCGTCAGACAGCTTCTAAGTGGGCTCATACCTGGAACAC 84004
QY 825 ATCTGGGGAAAGATATTAAGCATGAGTTCCTCAAAAAGAGCGCTGGCCATGAATAC 884
DB 84005 GTTCGGGGAGAGATATTAAGCATGATATATCCCAAAAAGAGCATGATGACATAC 84064
QY 885 CTACTGTATAGTGGCAGCAAAACCGAATGAGATCTCATTTGAAAAAAATATATGATGT 944
DB 84065 CTACTGTATAGTGGCAGCAAAACCGAATGAGATCTCATTTGAAAAAAATATATGATGT 84124
QY 945 ATAATGTGAT--TTTTTAAATAGAGGGGACCTTATTTATTTGTGTGTTAACTGATA 1002
DB 84125 ATTAATGTATTTGTTTTTAAAGAAATGGGGAAACTTTTATTTGTTAATGTTAACTGATA 84184
QY 1003 GGAATATGTACATCTATGTTTCATGATAGGTTGTTTCCCATTTAGCGAATGC 1062
DB 84185 GGAATATGTACATCTATTTTCATGATAGTGTGATTTTCTCACATTTAAGCAGGAATGC 84244
QY 1063 AATATAAAAATGTCGTTTTTTAA 1086
DB 1063 AATATAAAAATGTCGTTTTTTAA 1086

```









AC108660  
 LOCUS 323793 bp DNA linear HMG 22-SEP-2002  
 DEFINITION Rattus norvegicus clone CH230-74G23, \*\*\* SEQUENCING IN PROGRESS  
 AC108660  
 AC108660.4 GI:23101185  
 HG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
 Rattus norvegicus (Norway rat)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus

# REFERENCE AUTHORS

1 (bases 1 to 323793)  
 Muzny, D., Marie, Metzker, M., Lee, A., Adams, C., Alder, J., Allen, C., Allen, H., Alsebrook, S., Amin, A., Angiano, D., Anyalele, V., Ayagi, A., Ayodeji, M., Bacc, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Burch, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesaar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Din, H., Diya, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Georgergis, E., Geer, K., Galla, R., Garcia, A., Garner, T., Garza, M., Gharane, P., Haaland, M., Hamill, C., Hamilton, C., Hamilton, W., Harvey, Y., Havlak, P., Hayes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoque, M., Hollins, B., Howells, S., Hu, Y., Hume, J., Idler, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Kapathy, S., Kelly, S., Kelly, S., Khan, Z., Kiyav, C., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, Z., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, D., Lorenshewe, L., Louised, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Marzke, S., McLeod, M. P., McNeill, T. Z., Meenan, E., Miliashvili, A., Miner, G., Minja, E., Montemayor, J. J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nait, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokwelu, O., Okunolu, G., Olarinmoye, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., Plazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsberg, A., Sisson, I., Sitter, C. D., Sma, J., Speed, A., Sodergren, E., Song, X., Z., Sorrelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Xu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.

## JOURNAL

### COMMENT

Submitted (22-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On Sep 18, 2002 this sequence version replaced GI:21737904.  
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence contigs will be indicated in the feature table.

#### Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GRYA

Center clone name: CH230-74G23

Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 202774 bases at least Q40

Consensus quality: 205089 bases at least Q30

Estimated insert size: 226383; sum-of-contigs estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

NOTE: This is a 'working draft' sequence. It currently

consists of 6 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence.

as soon as it is available and the accession number will

be preserved.

1 312024: contig of 312024 bp in length

312025 312124: gap of unknown length

312125 316461: contig of 4337 bp in length

316462 316561: gap of unknown length

316562 317688: contig of 1127 bp in length

317689 317788: gap of unknown length

317789 319138: contig of 1350 bp in length

319139 319238: gap of unknown length

319239 321271: contig of 2033 bp in length

321272 321371: gap of unknown length

321372 323793: contig of 2422 bp in length.

Location/Qualifiers

1. 323793

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-74G23"

/complement(1136..1658)

/note="clone\_boundary"

clone\_end:77

site:ECORI

end\_sequence: BH37444"

## FEATURES

### source

#### misc-feature

#### misc-feature

#### misc-feature

BASE COUNT 65314 a 38957 c 39191 g 64287 t 116004 others  
 ORIGIN

Query Match 15.6%; Score 171.2; DB 2; length 323793;  
 Best Local Similarity 74.2%; Pred. No. 5.9e-27;  
 Matches 245; Conservative 0; Mismatches 78; Indels 7; Gaps 2;

```
OY 763 CTTCAAGCCTTATAGTTGCTGCAGACAGCTTCTAAGTGGGTCATACCTGGAACA 822
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 74209 CTTCAAGCCTTATAGTTGCTGCAGACAGCTTCAAACTGGGGTCCATACCTGAAAA 74268
OY 823 ACATCGTGGGGAAGATATAAGACATGTTCTTAAAAAGAGGCTGGCCATGAAT 882
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 74269 ACATCGTGGGAGAGATATAAGACATGCTGAACCTGCAAGAGAGACTGACCAAGAAAT 74328
OY 883 ACCTACTGTAGTGGCAGAGAAAACGGAAATGAGATCTCA---TTGAAAAAAATATATG 939
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 74329 ACCTACTGTAGTGGCAGAGAAAACGAATGAGATGTCACCTTTGAAAAATGGAAT 74388
OY 940 ATTGTATATGTGATTTTATAGAAATAGGGGACCTTATTTATTTGTGTACTGA 999
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 74389 GTTCATATATGTGATTTTGTAGAGTAGGGGGAATTTTAAATTTATTTATATCA 74448
OY 1000 ATAGGAAATGTACATACATATGTCATGATAGGGT---GATTTTTCCTTAAAGCA 1055
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 74449 CTAGGGGAGTGTACCTGTCAGTACAGGATGTAGTACTTTACCACTTAATCA 74508
OY 1056 GGAATGCATATPAAAAATGTGGTTTTTTT 1085
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 74509 AGAATGAATGTAAAAATGGAATTAATCA 74538
```

Search completed: August 27, 2003, 11:20:46  
job time : 4380 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 17:23:25 ; Search time 85 Seconds

(Without alignments)  
655,757 Million cell updates/sec

Title: US-09-940-227-83

Perfect score: 1198  
Sequence: 1 SEETITTTIDLPFKVKKM.....VPRKAGHEIPVSGSRKPE 216

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database :

1: SP\_ARCHAEA:\*  
2: SP\_BACTERIA:\*  
3: SP\_FUNGI:\*  
4: SP\_HUMAN:\*  
5: SP\_INVERTEBRATE:\*  
6: SP\_MAMMAL:\*  
7: SP\_MHC:\*  
8: SP\_ORGANELLE:\*  
9: SP\_PHAGE:\*  
10: SP\_PLANT:\*  
11: SP\_PROTOIST:\*  
12: SP\_VIRUS:\*  
13: SP\_VERTEBRATE:\*  
14: SP\_UNCLASSIFIED:\*  
15: SP\_VIRUS:\*  
16: SP\_BACTERIAP:\*  
17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1176	98.2	642	4	Q9UN76
2	1106.5	92.4	331	4	O15003
3	1074	89.6	638	11	Q9JMA9
4	1074	89.6	638	11	O91Y60
5	1074	89.6	638	11	O9A183
6	1066	89.0	638	11	O9D317
7	508.5	42.4	791	11	O91XQ2
8	508.5	42.4	799	11	O8CFM5
9	344.5	28.8	633	5	O8VC47
10	332	27.7	631	5	O9V7R0
11	332	27.7	631	5	O9NB97
12	327	27.3	619	11	O9J141
13	327	27.3	619	11	O8C009
14	327	27.3	629	13	O902V1
15	322	26.9	619	11	O9R0X6
16	319	26.6	620	6	O9GJ75

17	319	26.6	620	6	O9GJ76	O9GJ76 macaca fasc
18	317	26.5	620	6	O9GJ77	O9GJ77 salmrl scl
19	317	26.5	1201	5	O9WJ10	O9WJ10 drosophila
20	315	26.3	658	5	O94917	O94917 drosophila
21	315	26.3	943	5	O9VR07	O9VR07 drosophila
22	315	26.3	943	5	O02003	O02003 drosophila
23	311	26.0	674	5	O8WPM9	O8WPM9 olkopleura
24	307.5	25.7	599	4	O8N4K8	O8N4K8 homo sapien
25	305	25.5	635	13	O9DCN5	O9DCN5 gallus gall
26	303	25.3	630	13	O42482	O42482 rana catesb
27	302.5	25.3	629	13	O9DER6	O9DER6 oreochromis
28	300.5	25.1	628	13	O91503	O91503 torpeda mar
29	299.5	25.0	628	4	O96KH8	O96KH8 homo sapien
30	299	25.0	328	4	O9BW84	O9BW84 homo sapien
31	299	25.0	572	6	O9MY77	O9MY77 macaca mulle
32	298	25.0	617	6	O9MY78	O9MY78 macaca mulle
33	298	24.9	739	5	O9VJR4	O9VJR4 drosophila
34	298	24.9	639	5	O9NKA8	O9NKA8 drosophila
35	294.5	24.6	625	13	O9IBH6	O9IBH6 cyprinus ca
36	293.5	24.5	135	6	O28498	O28498 macaca mulle
37	293	24.5	617	11	O8R212	O8R212 mus musculu
38	292	24.4	567	11	O9WTR3	O9WTR3 rattus norv
39	292	24.4	597	11	O9WTR4	O9WTR4 rattus norv
40	292	24.4	598	13	O73771	O73771 raja sp. ga
41	292	24.4	617	11	O63380	O63380 rattus norv
42	290.5	24.2	625	13	O8JHX8	O8JHX8 salmo salar
43	290	24.2	374	5	O962P1	O962P1 aplysia cal
44	289	24.1	640	5	O8ISW0	O8ISW0 aedes aegypt
45	288	24.0	598	13	O91494	O91494 torpeda cal

## ALIGNMENTS

## RESULT 1

ID: Q9UN76

PREDIMINARY; PRT; 642 AA.

AC: Q9UN76;

DT: 01-MAY-2000 (TRENBLUREL.13, created)

DT: 01-MAY-2000 (TRENBLUREL.13, last sequence update)

DT: 01-OCT-2002 (TRENBLUREL.22, last annotation update)

DE: Amino acid transporter B0+.

GN: ATB0+.

OS: Homo sapiens (Human).

OC: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC: Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX: NCBI\_TaxID=9606;

RP: TISSUE=Mammary gland;

RC: MEDLINE=9376996; PubMed=10446133;

RA: Sloan J.L., Mager S.;

RT: "Cloning and functional expression of a human Na(+) and Cl(-)-

RL: J. Biol. Chem. 274:23740-23745(1999).

DR: EMBL; AF151978; AAD49223.1; -

DR: Genew; HGNC:11047; SLC6A14.

DR: InterPro; IPR00175; Na/ntran\_symport.

DR: Pfam; PF00209; SNF. 1.

DR: PRINTS; PR00176; NANEUSMPORT.

DR: PRODOM; PD000448; Na/ntran\_symport; 2.

DR: PROSITE; PS00610; NA\_NEUROTAN\_SYM\_1; 1.

DR: PROSITE; PS50267; NA\_NEUROTAN\_SYM\_3; 1.

SO: SEQUENCE 642 AA; 72152 MW; E0FCDD5F173128C0 CRC64;

Query Match 98.2%; Score 1176; DB 4; Length 642;

Best Local Similarity 98.6%; Pred. No. 1.7e-104;

Matches 213; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SEETITTTIDLPFKVKKRVPTTGCCVLFLGLGVCYAGIYVWHLIDHFCAGWG1 60

Db 427 SIEITTTTIDLPFKVKKRVPTTGCCVLFLGLGVCYAGIYVWHLIDHFCAGWG1 486

```

QY 61 LIAAILELGIIMWYGNRFIEDTEMMIGAKRMIFWLMWRACWFVITPILIAIFWLSLV 120
DB 487 LIAAILELGIIMWYGNRFIEDTEMMIGAKRMIFWLMWRACWFVITPILIAIFWLSLV 546
QY 121 QFHRPNYGAIPYDPMGVALGCMIVFCIIWIPIMAIKIIQAKGNIFORLISCRPASNM 180
DB 547 QFHRPNYGAIPYDPMGVALGCMIVFCIIWIPIMAIKIIQAKGNIFORLISCRPASNM 606
QY 181 GPLYEQRGERYKDMVVPKKEAGHEIPTVSGSKRPE 216
DB 607 GPLYEQRGERYKDMVVPKKEAGHEIPTVSGSKRPE 642

RESULT 2
ID 015003 PRELIMINARY; PRT; 331 AA.
AC 015003;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE GLYT-1 like (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Craftam-D
RA Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z96810; CAB09650.1;
DR InterPro: IPR001175; Na/ntrn_symport.
DR Pfam: PF00209; SNF. 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR PRODOM: PD000448; Na/ntrn_symport. 1.
DR PROSITE: PS50267; NA_NEUROTRAN_SYM_3; 1.
FT NON_TER 1
FT NON_TER 331
SQ SEQUENCE 331 AA; 37479 MW; 3B761927404F4E57 CRC64;

Query Match 92.4%; Score 1106.5; DB 4; Length 331;
Best Local Similarity 94.4%; Pred. No. 4; 1e-98;
Matches 204; Conservative 0; Mismatches 3; Indels 9; Gaps 2;

QY 1 SEETITTTTODLPPKVKMKRVPITLGGCLVFLGLGVCVTOAGIYWHLIDHFCAGWG1 60
DB 117 SITITTTTODLPPKVKMKRVPITLGGCLVFLGLGVCVTOAGIYWHLIDHFCAGWG1 175
QY 61 LIAAILELGIIMWYGNRFIEDTEMMIGAKRMIFWLMWRACWFVITPILIAIFWLSLV 112
DB 176 LIAAILELGIIMWYGNRFIEDTEMMIGAKRMIFWLMWRACWFVITPILIAIFWLSLV 235
QY 113 AITFWSLVQFHRPNYGAIPYDPMGVALGCMIVFCIIWIPIMAIKIIQAKGNIFORLIS 172
DB 236 AITFWSLVQFHRPNYGAIPYDPMGVALGCMIVFCIIWIPIMAIKIIQAKGNIFORLIS 295
QY 173 CCRPASNMGYLEQHRGERYKDMVVPKKEAGHEIPT 208
DB 296 CCRPASNMGYLEQHRGERYKDMVVPKKEAGHEIPT 331

RESULT 3
ID 09JMA9 PRELIMINARY; PRT; 638 AA.
AC 09JMA9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Colonc system B0+ amino acid transporter.
GN SIC6A14 OR MCATB0+.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ddy;
RA Ugua S.;
RT "Mouse Colonc System B0+ Amino Acid Transporter."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB033285; BAA94300.1;
DR MGD: MGI:1890216; Slc6a14.
DR InterPro: IPR000175; Na/ntrn_symport.
DR Pfam: PF00209; SNF. 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR PRODOM: PD000448; Na/ntrn_symport. 2.
DR PROSITE: PS00610; NA_NEUROTRAN_SYM_1; 1.
DR PROSITE: PS50267; NA_NEUROTRAN_SYM_3; 1.
SQ SEQUENCE 638 AA; 71455 MW; 5D5A78E2DFE7E612 CRC64;

Query Match 89.6%; Score 1074; DB 11; Length 638;
Best Local Similarity 86.1%; Pred. No. 1e-94;
Matches 186; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 SEETITTTTODLPPKVKMKRVPITLGGCLVFLGLGVCVTOAGIYWHLIDHFCAGWG1 60
DB 423 SITITTTTODLPPKVKMKRVPITLGGCLVFLGLGVCVTOAGIYWHLIDHFCAGWG1 482
QY 61 LIAAILELGIIMWYGNRFIEDTEMMIGAKRMIFWLMWRACWFVITPILIAIFWLSLV 120
DB 483 LIAAILELGIIMWYGNRFIEDTEMMIGAKRMIFWLMWRACWFVITPILIAIFWLSLV 542
QY 121 QFHRPNYGAIPYDPMGVALGCMIVFCIIWIPIMAIKIIQAKGNIFORLISCRPASNM 180
DB 543 KFHRRPYADIYDPMGVALGCMIVFCIIWIPIMAIKIIQAKGNIFORLISCRPASNM 602
QY 181 GPLYEQRGERYKDMVVPKKEAGHEIPTVSGSKRPE 216
DB 603 GPLYEQRGERYKDMVVPKKEAGHEIPTVSGSKRPE 638

RESULT 4
ID 091Y60 PRELIMINARY; PRT; 638 AA.
AC 091Y60;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Na+ and Cl-coupled neutral and basic amino acid transporter
DE ATB0+.
GN SIC6A14 OR ATB0.++.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=COLON;
RX MEDLINE=21203702; PubMed=11306607;
RA Hatanaka T., Nakashishi T., Huang W., Leibach F.H., Prasad P.D.,
RA Ganapathy V., Ganapathy M.E.;
RT "Na(+) and Cl(-)-coupled active transport of nitric oxide synthase
RT inhibitors via amino acid transport system B(0,+).";
RL J. Clin. Invest. 107:1035-1043(2001).
DR EMBL: AF320226; AAK43541.1;
DR MGD: MGI:1890216; Slc6a14.
DR InterPro: IPR000175; Na/ntrn_symport.
DR Pfam: PF00209; SNF. 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR PRODOM: PD000448; Na/ntrn_symport. 2.
DR PROSITE: PS00610; NA_NEUROTRAN_SYM_1; 1.
DR PROSITE: PS50267; NA_NEUROTRAN_SYM_3; 1.
SQ SEQUENCE 638 AA; 71428 MW; C68188E85F9837 CRC64;

Query Match 89.6%; Score 1074; DB 11; Length 638;
Best Local Similarity 86.1%; Pred. No. 1e-94;
Matches 186; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

```

```

OY 1 SEETITTTIODLPFKVKKRMVPTLGGCLVFLGLGVCYQAGIYVHLLDHFCAWGCI 60
DB 423 SEITTTTTFODLPFKAKRMVPTLGGCLLFLGLGCLVQAGIYVHLLDHFCAWGCI 482
OY 61 LIAAILELVGIWYIGNRFEDTEMIGAKRMIFWLMWRACWFYITPILLAIFWISLV 120
DB 483 LIAAILELVGIWYIGNRFEDTEMIGAKRMIFWLMWRACWFYITPILLAIFWISLV 542
OY 121 QFHRPNVGAIPYPMGVALGCMIVFCIIMIPMAIKIIOAKGNIFORLISCRPASNM 180
DB 543 KFHPRVDYADIPYPMGVALGCMIFCIIIMIPMAIKIIOAKGNIFORLISCRPASNM 602
OY 181 GPYLEOHNGERYKDMVVPKKEAGHEIPTVSGSRKE 216
DB 603 GPYLEKHNGERYKDMVPAKETDHEIPTISGSTKE 638

RESULT 5
O9R183 PRELIMINARY; PRT: 638 AA.
AC 09R183;
DB 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
GN Amino acid transporter B0+.
OS SLC6A14 OR ATB0+.
OC Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c; TISSUE=Lung;
RX MEDLINE=99376966; PubMed=10446133;
RA Sloan J.L., Mager S.;
RT "Cloning and functional expression of a human Na(+) and Cl(-)-
dependent neutral and cationic amino acid transporter B(0+).";
RT J. Biol. Chem. 274:23740-23745(1999).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c; TISSUE=Lung;
RA Reveill L., Sloan J.L., Mager S.;
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF161714; AAD49320.1; -.
DR MGD: MGI:1890216; SLC6A14.
DR InterPro: IPR000175; Na/nttran_sympor.
DR Pfam: PF00209; SNF; 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR PRODOM: PD000448; Na/nttran_sympor. 2.
DR PROSITE: PS00610; NA_NEUOTRAN_SYMP_1; 1.
DR PROSITE: PS50267; NA_NEUOTRAN_SYMP_3; 1.
SQ SEQUENCE 638 AA; 71332 MW; 5607420A394D5092 CRC64;

Query Match 89.6%; Score 1074; DB 11; Length 638;
Best Local Similarity 86.1%; Pred. No. 1e-94;
Matches 186; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

```

```

RESULT 6
O9D317 PRELIMINARY; PRT: 638 AA.
AC 09D317;
DB 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
GN 9030613J17Rik protein.
OS SLC6A14 OR 9030613J17Rik.
OC Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Salto T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Salto K.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Felschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakai R., Pesole G., Quackenbush J.,
RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Botfield D., Boujunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski B., Yoshida K., Hasegawa Y., Kawai H., Kontauki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
DR EMBL: AK018553; BAB31272.1; -.
DR MGD: MGI:1890216; SLC6A14.
DR InterPro: IPR00175; Na/nttran_sympor.
DR Pfam: PF00209; SNF; 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR PRODOM: PD000448; Na/nttran_sympor. 2.
DR PROSITE: PS00610; NA_NEUOTRAN_SYMP_1; 1.
DR PROSITE: PS50267; NA_NEUOTRAN_SYMP_3; 1.
SQ SEQUENCE 638 AA; 71412 MW; 5D461534187E7612 CRC64;

Query Match 89.0%; Score 1066; DB 11; Length 638;
Best Local Similarity 85.6%; Pred. No. 5.8e-94;
Matches 185; Conservative 14; Mismatches 17; Indels 0; Gaps 0;

```





ID Q9V7R0 PRELIMINARY; PRT; 631 AA.  
 AC Q9V7R0; Q961H9;  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
 DE CG8380 protein (GH22929p) (Dopamine transporter).  
 GN DAT OR CG8380.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Berns P.V., Berman B.P., Bhandari D., Bolintinas S.,  
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Buttle K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,  
 RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh N.V., Moberly C., Morris D., Moshrefi A.,  
 RA Merkulov G., Milshina N.V., Moberly C., Morris D., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC Gelniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Bisam D.A.,  
 RA Carlson J.M., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ilegwam C., Jaiswal M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,  
 RA Paclob J., Paragas V., Park S., Patel S., Pfeiffer B., Scheeler F.,  
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of Drosophila melanogaster genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.

RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.M., Celniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Milburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of Drosophila melanogaster genome.";  
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Flybase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guatin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nunco J., Paclob J., Paragas V., Park S., Phouenavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Gu H.H., Wu X.;  
 RT "Phenylalanine 105 in transmembrane domain 2 of the dopamine  
 transporter is involved in cocaine binding.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003806; AAF5986.2;  
 DR EMBL: AY051579; AAK93003.1;  
 DR EMBL: AF439752; AAL32055.1;  
 DR FLYBASE: FBgn004136; DAT.  
 DR InterPro: IPR000175; Na/ntan-symport.  
 DR PRINTS: PR00176; NANEUSMPORT.  
 DR PROSITE: PS00610; NA-NEUROTRAN\_SYM\_1; 1.  
 DR PROSITE: PSS0267; NA-NEUROTRAN\_SYM\_3; 1.  
 SO SEQUENCE 631 AA; 70330 MW; E7582D17F81E3166 CRC64;  
 Query Match 27.7%; Score 332; DB 5; Length 631;  
 Best Local Similarity 36.8%; Pred. No. 1,4e-23;  
 Matches 64; Conservative 20; Mismatches 77; Indels 4; Gaps 2;  
 QY 3 ETITTTTODLEPRKMKMKRPITIGCCLVLELVCTVQAGTYWHLIDHFCAGKGLI 62  
 DB 427 EAITTAISDEPKI-KNRRELFVAGLPSLVFVGLASCTGGFFHLIDRYAGSYILV 485  
 QY 63 AAILLEGIIMWYGNFIEDTEKMGAKRMIFLWMRACMFVTPITLLAIFWISLVQF 122  
 DB 486 AVFPEALAVMSITGNFSEDIRDMISFPP---GRYQVCNREVPALFLLFTYGLIGY 542  
 QY 123 HRPYGAIPYDGMGVALGMCWIVFCIIIPIMAIKIKIIOAKNIFQRLISCRP 176  
 DB 543 EPIITYADYVPSNMANALGMCWICASSVYVIMPAVIAFKLSIPGSLRGQFTITTP 596  
 RESULT 11  
 Q9NB97 PRELIMINARY; PRT; 631 AA.  
 AC Q9NB97;  
 DT 01-OCT-2000 (TRENBLREL. 15, Created)  
 DT 01-OCT-2002 (TRENBLREL. 15, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE Dopamine transporter.  
 GN DAT OR CG8380.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;

RN [11]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Head;  
 RA Porzgen P., Sonders M.S., Reed A.I., Ingram S.L., Amara S.G.;  
 RT Identification of neurotransmitter transporters from Drosophila  
 RL melanogaster";  
 RL Abstr. Soc. Neurosci. 25:160-160(1999).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Head;  
 RA Porzgen P., Park S.K., Hirsh J., Sonders M.S., Amara S.G.;  
 RT "An antidepressant-sensitive dopamine transporter from Drosophila  
 RL melanogaster";  
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF260833; AAF76882.1; -;  
 DR FlyBase: FBgn0034136; DAF;  
 DR InterPro: IPR000175; Na/ntran\_symport.  
 DR Pfam: PF00209; SNF; 1.  
 DR PRINTS: PR00176; NANEUSMPORT.  
 DR PRODOM: PD000448; Na/ntran\_symport; 1.  
 DR PROSITE: PS00610; NA\_NEUROTAN\_SYM\_1; 1.  
 DR PROSITE: PS0267; NA\_NEUROTAN\_SYM\_3; 1.  
 DR PROSITE: PS0267; NA\_NEUROTAN\_SYM\_3; 1.  
 SQ SEQUENCE 631 AA; 70349 MW; ABE3B53F5A91ED89 CRC64;

Query Match 27.7%; Score 332; DB 5; Length 631;  
 Best Local Similarity 36.8%; Pred. No. 1,4e-23;  
 Matches 64; Conservative 29; Mismatches 77; Indels 4; Gaps 2;

QY 3 ETTTITIDLPKVKMKRVPTLGGCLVLGLVCYQAGIYVWHLIDHFCAGMGLI 62  
 DB 427 EATITLSDPEFKT-KRNRLEFLVAGLFLYFVVGGLASCTQGGFPHLDRAAGSYILV 485  
 QY 63 AATLELVGIIWYGGRTFEDTEMIGAKRWIFLWRCQWPIPTLLIAIFWISLYOF 122  
 DB 486 AATLELVGIIWYGGRTFEDTEMIGAKRWIFLWRCQWPIPTLLIAIFWISLYOF 542  
 QY 123 HRPNYGAIYPPDWGVALGCMIVFCIIWIPIMAIKIIQAKNIFQRLISCRP 176  
 DB 543 EPLTVADYVYPSMANLGMCIAGSSVMIIPAVAFILSLTPGSLRQRTILTPP 596

RESULT 12  
 Q9J41 PRELIMINARY; PRT; 619 AA.  
 AC 09J41;  
 DT 01-OCT-2000 (Tremblrel. 15, Created).  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Dopamine transporter.  
 GN SIC6A3 OR DAT.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RA Hein A., Ortmann R., Mueller K., Luebbert H., Frenzel S.;  
 RT "A rapid and simple flow cytometric method to investigate apoptosis  
 RL and oncosis in adherent cells";  
 RL Submitted (Nov-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF109391; AAF85795.1; -;  
 DR MGI: MGI:94862; SIC6A3.  
 DR InterPro: IPR000175; Na/ntran\_symport.  
 DR Pfam: PF00209; SNF; 1.  
 DR PRINTS: PR00176; NANEUSMPORT.  
 DR PRODOM: PD000448; Na/ntran\_symport; 1.  
 DR PROSITE: PS00610; NA\_NEUROTAN\_SYM\_1; 1.  
 DR PROSITE: PS00754; NA\_NEUROTAN\_SYM\_2; 1.  
 DR PROSITE: PS0267; NA\_NEUROTAN\_SYM\_3; 1.  
 SQ SEQUENCE 619 AA; 68819 MW; ELD6137BCFA61E9B CRC64;

Query Match 27.3%; Score 327; DB 11; Length 619;

Best Local Similarity 36.8%; Pred. No. 4e-23;  
 Matches 64; Conservative 32; Mismatches 74; Indels 4; Gaps 2;

QY 3 ETTTITIDLPKVKMKRVPTLGGCLVLGLVCYQAGIYVWHLIDHFCAGMGLI 62  
 DB 427 EATITLSDPEFKT-KRNRLEFLVAGLFLYFVVGGLASCTQGGFPHLDRAAGSYILV 485  
 QY 63 AATLELVGIIWYGGRTFEDTEMIGAKRWIFLWRCQWPIPTLLIAIFWISLYOF 122  
 DB 486 AATLELVGIIWYGGRTFEDTEMIGAKRWIFLWRCQWPIPTLLIAIFWISLYOF 542  
 QY 123 HRPNYGAIYPPDWGVALGCMIVFCIIWIPIMAIKIIQAKNIFQRLISCRP 176  
 DB 543 EPLTVADYVYPSMANLGMCIAGSSVMIIPAVAFILSLTPGSLRQRTILTPP 596

RESULT 13  
 Q8C009 PRELIMINARY; PRT; 619 AA.  
 AC 08C009;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Dopamine transporter.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Olfactory brain;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium.  
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT Analysis of the mouse transcriptome based on functional annotation of  
 RL 60,770 full-length cDNAs.;  
 RL Nature 420:563-573(2002).  
 DR EMBL: AC032629; BAC27959.1; -;  
 SQ SEQUENCE 619 AA; 68888 MW; 19F446221807535D CRC64;

Query Match 27.3%; Score 327; DB 11; Length 619;  
 Best Local Similarity 36.8%; Pred. No. 4e-23;  
 Matches 64; Conservative 32; Mismatches 74; Indels 4; Gaps 2;

QY 3 ETTTITIDLPKVKMKRVPTLGGCLVLGLVCYQAGIYVWHLIDHFCAGMGLI 62  
 DB 427 EATITLSDPEFKT-KRNRLEFLVAGLFLYFVVGGLASCTQGGFPHLDRAAGSYILV 485  
 QY 63 AATLELVGIIWYGGRTFEDTEMIGAKRWIFLWRCQWPIPTLLIAIFWISLYOF 122  
 DB 486 AATLELVGIIWYGGRTFEDTEMIGAKRWIFLWRCQWPIPTLLIAIFWISLYOF 542  
 QY 123 HRPNYGAIYPPDWGVALGCMIVFCIIWIPIMAIKIIQAKNIFQRLISCRP 176  
 DB 543 EPLTVADYVYPSMANLGMCIAGSSVMIIPAVAFILSLTPGSLRQRTILTPP 596

RESULT 14  
 Q90ZV1 PRELIMINARY; PRT; 629 AA.  
 AC 090ZV1;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Dopamine transporter.  
 GN SIC6A3.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE-21152910; PubMed-11231083;  
 RA Holtschuh J., Ryu S., Aberger F., Driever W.,  
 RT "Dopamine transporter expression distinguishes dopaminergic neurons  
 RT from other catecholaminergic neurons in the developing zebrafish  
 RT embryo.";  
 RL Mech. Dev. 101:237-243(2001).  
 DR EMBL: AF318177; AAK52449.1; -  
 DR ZFIN: ZDB-GENE-010316-1; slc6a3.  
 DR InterPro: IPR000175; Na/ntan\_symport.  
 DR Pfam: PF00209; SNF; 1.  
 DR PRINTS: PR00176; NANEUSMPORT.  
 DR PRODOM: PD000448; Na/ntan\_symport; 1.  
 DR PROSITE: PS00610; NA\_NEUROTRAN\_SYMP\_1; 1.  
 DR PROSITE: PS00754; NA\_NEUROTRAN\_SYMP\_2; 1.  
 DR PROSITE: PS50267; NA\_NEUROTRAN\_SYMP\_3; 1.  
 SQ SEQUENCE 629 AA; 70161 MW; EF7CD46C3078C8CB CRC64;

Query Match 27.3%; Score 327; DB 13; Length 629;  
 Best Local Similarity 34.9%; Pred. No. 4.1e-23;  
 Matches 66; Conservative 33; Mismatches 80; Indels 10; Gaps 3;

OY 3 ETITTTIODELPKVMKKMRVPTLTGCVLFLGLVCVYQAGIYVWHLIDHFCAGWGILI 62  
 Db 437 ESVITGLIDEF-KFLKHKRELFTLTIVSTFLSLICVINGIYVFTLLDHPAAGTSLIF 495  
 OY 63 AALELVGIITWYGNRFEDTEMMIGAKRMIFLMPACWFTVTPILLIAIFISLVQF 122  
 Db 496 GVLIEAIGVAMFYGVDFSDIDIEIMIGRP--GLYWRLCWKVSPCFLEVVVVSFATF 552  
 OY 123 HRPNYGAIYPDPMGVAGLWCMIVFCIWIPIPIAIKIIOAKNIFORLISCCRPASNMP 182  
 Db 553 NPKYGSYFFPTWAMWVGCISIMINVLPAFYKFCSLPSCDKLAVAITPET---- 608  
 OY 183 YLEQRGER 191  
 Db 609 --DHLVER 615

RESULT 15  
 O9ROX6 PRELIMINARY: PRT; 619 AA.  
 AC O9ROX6:  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Dopamine transporter.  
 GN SLG6A3 OR DAT1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BAIB/C; TISSUE=Brain;  
 RA Brues M., Wieland A., Bonisch H.;  
 RT "Molecular cloning and functional expression of the mouse dopamine  
 RT transporter.";  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ238309; CAB51926.1; -  
 DR MGI: 94862; SLG6A3  
 DR InterPro: IPR000175; Na/ntan\_symport.  
 DR Pfam: PF00209; SNF; 1.  
 DR PRINTS: PR00176; NANEUSMPORT.  
 DR PRODOM: PD000448; Na/ntan\_symport; 1.  
 DR PROSITE: PS00610; NA\_NEUROTRAN\_SYMP\_1; 1.  
 DR PROSITE: PS00754; NA\_NEUROTRAN\_SYMP\_2; 1.  
 DR PROSITE: PS50267; NA\_NEUROTRAN\_SYMP\_3; 1.  
 SQ SEQUENCE 619 AA; 68775 MW; F42BA5A959CC3FDE CRC64;

Query Match 26.9%; Score 322; DB 11; Length 619;  
 Best Local Similarity 36.2%; Pred. No. 1.2e-22;  
 Matches 63; Conservative 32; Mismatches 75; Indels 4; Gaps 2;

OY 3 ETITTTIODELPKVMKKMRVPTLTGCVLFLGLVCVYQAGIYVWHLIDHFCAGWGILI 62  
 Db 427 ESVITGLIDEF-QLKHKRELFTLTIVATFLSLICVANGIYVFTLLDHPAAGTSLIF 485  
 OY 63 AALELVGIITWYGNRFEDTEMMIGAKRMIFLMPACWFTVTPILLIAIFISLVQF 122  
 Db 486 GVLIEAIGVAMFYGVQFSDIKQWVGORP--NLVWRLCWKVSPCFLEVVVVSIVTF 542  
 OY 123 HRPNYGAIYPDPMGVAGLWCMIVFCIWIPIPIAIKIIOAKNIFORLISCCRP 176  
 Db 543 RPHYGATIFPDWAMALGMIITATSSMAVPIYATYKFCSLPSCDKLAVAITP 596

Search completed: August 27, 2003, 17:36:25  
 Job time : 90 secs

**THIS PAGE BLANK (USPTO)**

---

PT New Lung specific genes and their encoded proteins, useful in gene

inertias

PT therapy or as a vaccine for treating lung cancer, as well as for  
PT measuring metastases of lung cancer, or staging, monitoring, diagnosing  
PT or imaging lung cancer -  
XX  
PS Claim 1; Page 172; 206pp; English.

CC The invention describes a new lung specific gene and it's variants. The  
CC lung specific gene proteins and genes are useful in gene therapy or as a  
CC vaccine for treating lung cancer. Lung specific genes are also useful for  
CC staging, monitoring, diagnosing or imaging lung cancer, as well as for  
CC measuring metastases of lung cancer. This sequence represents a lung  
CC specific gene described in the invention.

Sequence 1094 BP; 338 A; 172 C; 236 G; 348 T; 0 other;

Query Match	100.0%	Score 1094;	DB 24;	Length 1094;
Best Local Similarity	100.0%;	Pred. No. 7.9e-270;		
Matches 1094; Conservative	0;	Mismatches	0;	Gaps 0;

QY 1 GTTCACAGGGGACGTCTACCTTACAGTTGTTATGATGAATAATCATATTAAGCAGACTA 60

Db 1 GTTACACAGGGGACTGTTACCTTACAGTTGTTATGATGAATAATCATATTAAGCAGACTA 60

QY	61	AAAAAGCAAAATAGCCGCTGCATTTTGATTTTCATCAGAACCAAGTGGTGCACGCCAAAAGA	120
Db	61	AAAAAGCAAAATAGCCGCTGCATTTTGATTTTCATCAGAACCAAGTGGTGCACGCCAAAAGA	120

QY	121	AAAAACCTGCTGCTTTTAAAGCAAGAACGGGCTCAGAGTCTATTTTCAACATATA	180
Db	121	AAAAACCTGCTGCTTTTAAAGCAAGAACGGGCTCAGAGTCTATTTTCAACATATA	180
	121	AAAAACCTGCTGCTTTTAAAGCAAGAACGGGCTCAGAGTCTATTTTCAACATATA	180

QY 181 TTGTAAATTGATTAAACATACATTGTAATTCGTGATTAAACAGTAAATGGCAATAAAGTTG 240

Db 181 TTGTAAATTGATTAAACATACATTGTAATTCGTGATTAAACAGTAAATGGCAATAAAGTTG 240

Qy	241	ACATATACATGTCATTTCTCCCCCTGTGAAGAAACGATCAACAACAATTTCAAGATT	300
Db	241	ACATATACATGTCATTTCTCCCCCTGTGAAGAAACGATCAACAACAATTTCAAGATT	300

QY	301	ATTTCCCAAAGTGATGAAGAATAATAGGGTTCCCATPACTTTGGCGCTGCTGTGGTTTT	360
Db	301	ATTTCCCAAAGTGATGAAGAATAATAGGGTTCCCATPACTTTGGCGCTGCTGTGGTTTT	360

Oy	361	GTTCCTCCCTGGTCTCGTGTGTGA	CTCAGGCTGGAATTTACTGGGTTCA	CTGATGTA	420
Db	361	GTTCCTCCCTGGTCTCGTGTGTGA	CTCAGGCTGGAATTTACTGGGTTCA	CTGATGTA	420

OY		421	CCACATCTCTGCTGGATGGGCGATTTTAATGCAGCCTATACTGAGCTAGTGTGAATCAT	480
D6		421	CCACTCTCTGCTGGATGGGCGATTTTAATGCAGCCTATACTGAGCTAGTGTGAATCAT	480

Qy 481 CTGAGTTTATGAGGGAACAGATTATTGAGCATACAGAAATGATGATTTGGAGCAAAAG 540

Db 481 CTGAGTTTATGAGGGAACAGATTATTGAGCATACAGAAATGATGATTTGGAGCAAAAG 540

QY 541 GTGATATTCCTGGCATGTGGAGAGCTTCTGTGTGTAATTACGCTATCTTTGAT 600

Db 541 GTGATATTCCTGGCATGTGGAGAGCTTCTGTGTGTAATTACGCTATCTTTGAT 600

QY	601	TGCATATTATTCGTGCATTTGCATTCATAGACCAATATTATGGCGCAATTCATA	660
Db	601	TGCATATTATTCGTGCATTTGCATTCATAGACCAATATTATGGCGCAATTCATA	660

Oy	661	CCCTGACGTGGGAGTTCCTTAGCGCGTGATGATTGTTCTGCATTATTGGATTCC	720
Dh	661	CCCCGACGTGGGGACCTGCTTTAAGGTCGCTCATATGATGCTTTCTGCAATTAATGGGATGCC	720

QY	721	AATTATGGCATCAAAAAATATTCAGGCTAAAGGAACATCTTCAACGGCTTAAAG	780
DB	721	AATTATGGCTATCAAAAAATATTCAGGCTAAAGGAACATCTTCAACGGCTTAAAG	780

781 TTGTCGACAGCAGCTTCTAAGTGGGTCCTACCTGGAAACAACATCGTGGGAAAGATA 840  
 |||

Db	781	TTGCTGCAGACACAGCTTCTA	CTGCGGGCCCACTACCTG	GAACAACTCGTGGGAAAAAGTA	840	
QY	841	TAAGACATG	TAGTTCCCTAAAAA	GAGCGCTGGCCATGA	AAATACCTACTGTTAGTGGCAG	900
Db	841	TAAGACATG	TAGTTCCCTAAAAA	GAGCGCTGGCCATGA	AAATACCTACTGTTAGTGGCAG	900
QY	901	CAGAAACCGGAATG	AGATGTCATTG	GA	AAAAATATATGATGATATGATGATTTT	960
Db	901	CAGAAACCGGAATG	AGATGTCATTG	GA	AAAAATATATGATGATATGATGATTTT	960
QY	961	TAGAAATAGGGGAC	CCATTATTTATTTGTG	GTAACTGA	TAAGAAAAATGTACATCTAT	1020
Db	961	TAGAAATAGGGGAC	CCATTATTTATTTGTG	GTAACTGA	TAAGAAAAATGTACATCTAT	1020
QY	1021	GTCATGATGATGGT	GTATTTTCCCATTTA	AGCAGAAATCA	ATATAAAAATGTGGTT	1080
Db	1021	GTCATGATGATGGT	GTATTTTCCCATTTA	AGCAGAAATCA	ATATAAAAATGTGGTT	1080
QY	1081	TTTTAAAAAAA	AAAA	1094	TTTTAAAAAAA	1094
Db	1081	TTTTAAAAAAA	AAAA	1094	TTTTAAAAAAA	1094

RESULT 2	—
ABV25142	—
ID	ABV25142 standard; cDNA; 4817 BP.

AA	ABV25142;	—
AC		
XX		
DT	16-SEP-2002	(first entry)

AA	DE	XX	KM
Human prostate expression marker	CDNA 25133		
Human; prostate cancer; cytostatic; carcinoma			

rw	pharmacogenomic market; gene; ss.
xx	
os	
xx	
	Homo sapiens.

PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX

PF	20-FEB-2001; 2001WO-US051/1
XX	
PR	17-FEB-2000; 2000US-183319P
PR	16-MAR-2000; 2000US-189862P

PR 25-MAY-2000; 2000US-20/454P;  
PR 09-JUN-2000; 2000US-211314P;  
PR 18-JUL-2000; 2000US-219007P;  
PR 13-DEC-2000; 2000US-255281P;

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC  
PA  
XX  
PI Schlegel R. Endege WO. Monahan JE:

XX		—
DR	WPI; 2001-662795/76.	
XX		—
PT	Novel isolated nucleic acid molecule	

PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
XX  
Claim 1: Page 4903-4904: 11750pm English

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement (I') is useful for:  
CC

CC (a) assessing whether a patient is afflicted with prostate cancer  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient.

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient.

CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or incidence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 4817 BP; 1430 A; 807 C; 963 G; 1617 T; 0 other;

Query Match 73.2% Score 800.6; DB 23; Length 4817;  
Best Local Similarity 98.9%; Pred. No. 1.5e-194;  
Matches 806; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
QY 272 GAAACGATCACAACACAAATTCAGATTATTTCCCAAGATGATGAGAAATAGAGGTT 331
D 1392 GAAACGATCACAACACAAATTCAGATTATTTCCCAAGATGATGAGAAATAGAGGTT 1451
QY 332 CCCATTAATCTGGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 391
D 1452 CCCATTAATCTGGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1511
QY 392 GCTGGAATTTACTGGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 451
D 1512 GCTGGAATTTACTGGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1571
QY 452 GCAGCTATTTACTGGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 511
D 1572 GCAGCTATTTACTGGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1631
QY 512 GATACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 571
D 1632 GATACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1691
QY 572 TGGTTGTAATTTACGCTATCTCTTTGATGCAATATTTACTGATGATGATGATGATGAT 631
D 1692 TGGTTGTAATTTACGCTATCTCTTTGATGCAATATTTACTGATGATGATGATGATGAT 1751
QY 632 CATAGACCTAATTTATGCGCAATTCATACCTGATGCGGAGTGGCTTATGAGCTGGTGT 691
D 1752 CATAGACCTAATTTATGCGCAATTCATACCTGATGCGGAGTGGCTTATGAGCTGGTGT 1811
QY 692 ATGATGTTTCTGCAATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 751
D 1812 ATGATGTTTCTGCAATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1871
QY 752 AAAGAAACATCTTTCAACGCTTATGATGATGATGATGATGATGATGATGATGATGATG 811
D 1872 AAAGAAACATCTTTCAACGCTTATGATGATGATGATGATGATGATGATGATGATGATG 1931
QY 812 TACCTGGACACATCTGCGGGAAGATATTAAGACATGCTGATGATGATGATGATGATG 871
D 1932 TACCTGGACACATCTGCGGGAAGATATTAAGACATGCTGATGATGATGATGATGATG 1991
QY 872 GGCATGAAATTAATCTGTTAGTGGAGAGAAACCGGATGATGATGATGATGATGATG 931
D 1992 GGCATGAAATTAATCTGTTAGTGGAGAGAAACCGGATGATGATGATGATGATGATG 2051
QY 932 AATATATGATTTGATATGATTTTATGATGATGATGATGATGATGATGATGATGATGAT 991
D 2052 AATATATGATTTGATATGATTTTATGATGATGATGATGATGATGATGATGATGATGAT 2111
QY 992 TTAACGATATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1051
D 2112 TTAACGATATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2171
QY 1052 AGCAGAAATGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1086
D 2172 AGCAGAAATGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2206
```

RESULT 3  
AAH98176  
ID AAH98176 standard; cDNA; 2098 BP.  
XX

```
AC AAH98176;
XX 12-OCT-2001 (first entry)
DT
DE
XX
DE
XX
KW Human EST-derived coding sequence SEQ ID NO: 33.
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition; ss.
XX
OS Homo sapiens.
XX
XX WO200154477-A2.
XX
XX 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-US02687.
XX
XX 25-JAN-2000; 2000US-0491404.
XX 17-JUL-2000; 2000US-0617746.
XX 03-AUG-2000; 2000US-0631451.
XX 15-SEP-2000; 2000US-0663870.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
XX Cao Y, Drmanac RA, Zhang J, Werhman T;
XX
XX WPI: 2001-476164/51.
XX P-PSDB: AAM23517.
XX
XX Isolated polypeptide for treatment of diseases, diagnostics, raising
XX antibodies and research use.
XX
XX Claim 1; Page 205; 1275pp; English.
XX
XX The present invention provides the protein and coding sequences of novel
XX proteins from a variety of organisms, including human, dog, cat, horse,
XX cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
XX urchin and tomato. These were derived from expressed sequence tags (ESTs)
XX from the organism of interest. They can be used in diagnostics,
XX forensics, gene mapping, identification of mutations, to assess
XX biodiversity and for nutritional purposes. The present sequence is a cDNA
XX of the invention.
XX
XX Sequence 2098 BP; 581 A; 359 C; 469 G; 689 T; 0 other;

Query Match 72.9%; Score 798; DB 22; Length 2098;  
Best Local Similarity 98.8%; Pred. No. 5.1e-194;  
Matches 804; Conservative 0; Mismatches 10; Indels 0; Gaps 0;



```
QY 272 GAAACGATCACAACACAAATTCAGATTATTTCCCAAGATGATGAGAAATAGAGGTT 331
D 1285 GAAACGATCACAACACAAATTCAGATTATTTCCCAAGATGATGAGAAATAGAGGTT 1344
QY 332 CCCATTAATCTGGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 391
D 1345 CCCATTAATCTGGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1404
QY 392 GCTGGAATTTACTGGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 451
D 1405 GCTGGAATTTACTGGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1464
QY 452 GCAGCTATTTACTGGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 511
D 1465 GCAGCTATTTACTGGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1524
QY 512 GATACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 571
D 1525 GATACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1584
QY 572 TGGTTGTAATTTACGCTATCTCTTTGATGCAATATTTACTGATGATGATGATGATGAT 631
```


```





Human GLYTLIKE protein coding sequence #2.

xx GLYTLIKE protein: Glycine transporter protein; human; neuropathic pain;  
xx chronic pain; postoperative pain; rheumatoid arthritic pain; neuralgia;  
xx neuropathy; algesia; nerve injury; ischemia; neurodegeneration; stroke;  
xx inflammation; inflammatory disorder; spasticity; myoclonus; epilepsy;  
xx head trauma; chromosome localisation study; therapy; ss.  
xx  
xx Homo sapiens.  
xx  
xx WO200014221-A1.  
xx  
xx 16-MAR-2000.  
xx  
xx 03-SEP-1999; 99WO-GB02909.  
xx  
xx 04-SEP-1998; 98GB-0019405.  
xx  
xx (SMIK ) SMITHKLINE BEECHAM PLC.  
xx  
xx Duckworth DM, Evans JR:  
xx  
xx WPI: 2000-256973/22.  
xx  
xx P-PSDB; AAY81944.  
xx  
xx New human glycine transporter, known as GLYTLIKE, useful in the  
xx treatment of diseases such as neuropathic pain, chronic pain,  
xx neuralgia, ischemia, stroke, incontinence, spasticity and epilepsy,  
xx comprises a 642 amino acid sequence -  
xx  
xx  
xx Claim 11: Page 18; 33pp: English.  
xx  
xx This sequence encodes a human glycine transporter (GLYT) protein,  
xx designated GLYTLIKE. The GLYTLIKE polypeptide and the polynucleotide  
xx encoding it are useful in the treatment of diseases such as neuropathic  
xx pain, pain, chronic pain, postoperative pain, rheumatoid arthritic pain,  
xx neuralgia, neuropathies, algesia, nerve injury, ischemia,  
xx neurodegeneration, stroke, incontinence, inflammatory disorders,  
xx spasticity, myoclonus, epilepsy and head trauma. The DNA sequence is  
xx valuable for chromosome localisation studies. The GLYTLIKE polypeptide  
xx and polynucleotide are also useful in diagnostic assays for detecting  
xx diseases associated with in appropriate GLYTLIKE activity or levels.  
xx  
xx  
xx Sequence 2177 BP; 581 A; 408 C; 520 G; 668 T; 0 other;  
xx  
xx  
xx Query Match 57.9%; Score 633.8; DB 21; Length 2177;  
xx Best Local Similarity 98.9%; Pred. No. 5.1e-152;  
xx Matches 638; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 272 GAAACGATCACAACACAAATTCAGATTATTTCCAAAGATGAGAAATGAGGTT 331  
Db 1533 GAAACGATCACAACACAAATTCAGATTATTTCCAAAGATGAGAAATGAGGTT 1592

QY 332 CCCATTAACCTTTGGGCTGCTGCTTGGTTTCTCCTGGTCTGCTGTGACTCAG 391  
Db 1593 CCCATTAACCTTTGGGCTGCTGCTTGGTTTCTCCTGGTCTGCTGTGACTCAG 1652

QY 392 GCTGGAATTTACCTGGGTTTCATCTGATGACACCTTGTGCTGATGGGCAATTTAAT 451  
Db 1653 GCTGGAATTTACCTGGGTTTCATCTGATGACACCTTGTGCTGATGGGCAATTTAAT 1712

QY 452 GCACGCTTACTGAGCTAGTTGGATATCATCTGATTTATGGAGGAGACAGATTCATTAG 511  
Db 1713 GCACGCTTACTGAGCTAGTTGGATATCATCTGATTTATGGAGGAGACAGATTCATTAG 1772

QY 512 GATACAGAAATGATGATTGGAGCAAAAGAGTGTATTTCTGGCTATGSGTGAGAGCTTGC 571  
Db 1773 GATACAGAAATGATGATTGGAGCAAAAGAGTGTATTTCTGGCTATGSGTGAGAGCTTGC 1832

QY 572 TGGTTTGAATTTACGCCATCCCTTTTGATTTGCATATTTATTCGCTGATGSGTCATTTT 631  
Db 1833 TGGTTTGAATTTACGCCATCCCTTTTGATTTGCATATTTATTCGCTGATGSGTCATTTT 1892

QY	632	CATGACCTAATTTAAATGGGGCAATTCGATACCTGACCTGAGGAGGATGGCTTAAGCTGGT	691		
Db	1893	CATGACCTAATTTTGGCGCAATTCGATACCTGACCTGAGGAGGATGGCTTAAGCTGGT	1952		
QY	692	ATGATTTGTTTCTGCATTAATTTGGATTCGATTAATGCTATCAATAAAATTAATTCAGGCT	751		
Db	1953	ATGATTTGTTTCTGCATTAATTTGGATTCGATTAATGCTATCAATAAAATTAATTCAGGCT	2012		
QY	752	AAAGGAAACATCTTTCAACGCTTATTAAGTTGCTGCAGACCAAGCTTTACTGGGTCGA	811		
Db	2013	AAAGGAAACATCTTTCAACGCTTATTAAGTTGCTGCAGACCAAGCTTTACTGGGTCGA	2072		
QY	812	TACCTGGAAACAATCGTGGGGAAGATTAATAACAATGTAAGTCTTAAAAAAAGAGGCT	871		
Db	2073	TACCTGGAAACAATCGTGGGGAAGATTAATAACAATGTAAGTCTTAAAAAAAGAGGCT	2132		
QY	872	GGCCATGAAATACCTACTGTTAGTGGCAGCAGAAACCGGAAATGA	916		
Db	2133	GACCATGAAATACCTACTGTTAGTGGCAGCAGAAACCGGAAATGA	2177		
RESULT 6					
AAH30720	AAH30720 standard; cDNA; 382 BP.				
XX	AAH30720:				
XX	27-JUL-2001	(first entry)			
DT					
XX	Human colon cancer cell line Km12L4-A cDNA library derived sequence #654.				
DE					
XX	Human: diagnosis: colon cancer; cancer; malignant; chromosome mapping;				
KW	detection: colon cancer cell line Km12L4-A; ss.				
KN					
XX					
OS	Homo sapiens.				
XX					
FN	WO200018916-A2.				
PD	06-APR-2000.				
XX					
XX	23-SEP-1999;	99WO-US222226.			
PF					
PR	28-SEP-1998;	98US-0102161.			
PR	28-SEP-1998;	98US-0102180.			
PR	29-SEP-1998;	98US-0102380.			
PR	08-OCT-1998;	98US-0103815.			
PR	27-OCT-1998;	98US-0105877.			
XX					
PA	(CHIR ) CHIRON CORP.				
PA	(HYSE-) HXSEQ INC.				
XX					
PI	Williams LT, Escobedo J, Inmls MA, Garcia PD, Sudduth-Klinger J;				
PI	Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;				
PI	Lanson G, Dmanac R, Crivenjakov R, Dickson M, Dmanac S, Labot I;				
PI	Leshkowitz D, Kita D, Garcia V, Jones LM, Stache-Crain B;				
XX					
DR	WPI: 2000-293155/25.				
XX					
PT	Polynucleotide library comprising 1079 defined sequences, useful in				
PT	the form of an array to detect cancer or susceptibility to cancer -				
XX					
PS	Claim 1; Page 372; 502pp; English.				
XX					
CC	The present invention describes a library of polynucleotides comprising				
CC	1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described				
CC	are: (1) an isolated polynucleotide (I) having at least 90% identity to				
CC	one of the 1079 sequences; (2) a recombinant host cell containing (1);				
CC	(3) an isolated polypeptide (II) encoded by (1); (4) an antibody that				
CC	specifically binds to (II); (5) a vector comprising (I); and (6) a method				
CC	of detecting differentially expressed genes correlated with a cancerous				
CC	state of a mammalian cell comprising detecting a gene product encoded by				
CC	65 of the 1079 sequences given in the specification. The polynucleotides				
CC	are used to monitor patients having (or susceptible) to cancer to detect				

CC potentially malignant events at a molecular level before they are  
CC detectable at a gross morphological level. The polynucleotides are also  
CC useful for monitoring the efficacy of various therapies and preventive  
CC interventions. Polynucleotide probes based on the disclosed sequences  
CC are useful for chromosome mapping and detection of transcription levels.  
CC The 1079 polynucleotide sequences were derived from a human colon cancer  
CC cell line Kml214-A cDNA library.

XX  
SQ Sequence 382 BP; 89 A; 68 C; 92 G; 132 T; 1 other;

Query Match 25.6%; Score 280; DB 21; Length 382;  
Best local Similarity 99.6%; Pred. No. 8.8e-62;  
Matches 280; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 272 GAAGAGTACACAAACAATTCACAGTTTATTTCCCAAGTGATGAAGAAATGAGGGTT 331  
DB 96 GAACAGTACACAAACAATTCACAGTTTATTTCCCAAGTGATGAAGAAATGAGGGTT 155  
QY 332 CCCATTAACCTTGGGCGTGGCTGTTGTTTCTCCTGGTCTGCTGTGACTCAG 391  
DB 156 CCCATTAACCTTGGGCGTGGCTGTTGTTTCTCCTGGTCTGCTGTGACTCAG 215  
QY 392 GCTGGAATTTACTGGGTCATCTGATTTGACACCTTCTGCTGATGGGCAATTTAAT 451  
DB 216 GCTGGAATTTACTGGGTCATCTGATTTGACACCTTCTGCTGATGGGCAATTTAAT 275  
QY 452 GCAGCATATCTGAGCTACTTGAATCATCTGATTTATGAGGGAACAGATTCATGAG 511  
DB 276 GCAGCATATCTGAGCTACTTGAATCATCTGATTTATGAGGGAACAGATTCATGAG 335  
QY 512 GATACAGAAATGATGATTTGAGCAGCAAGAGGTGATTTCTG 552  
DB 336 GATACAGAAATGATGATTTGAGCAGCAAGAGGTGATTTCTG 376

RESULT 7  
AAA91843  
ID AAA91843 standard; cDNA; 2394 BP.  
XX  
AC AAA91843;  
DT 23-FEB-2001 (first entry)  
XX  
DE Human glycine transporter type 2 protein mutant coding sequence #3.  
XX  
KW Human glycine transporter type 2; GLYT2; muscle spasticity; myoclonus;  
KW epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;  
KW dystonia; Huntington's disease; amyotrophic lateral sclerosis;  
KW mutant; muten; ss.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2394  
FT /tag- a  
FT /product- "mutant Glyt2 protein"  
FT misc\_difference 70  
FT /tag- b  
FT /note- "wild-type G substituted by A"  
FT misc\_difference 77  
FT /tag- c  
FT /note- "wild-type C substituted by T"  
FT misc\_difference 220  
FT /tag- d  
FT /note- "wild-type C substituted by T"  
FT misc\_difference 244  
FT /tag- e  
FT /note- "wild-type C substituted by T"  
FT misc\_difference 266  
FT /tag- f  
FT /note- "wild-type C substituted by T"  
FT misc\_difference 304

FT /tag- g  
FT /note- "wild-type A substituted by G"  
FT misc\_difference 463  
FT /tag- h  
FT /note- "wild-type A substituted by G"  
FT misc\_difference 521  
FT /tag- i  
FT /note- "wild-type T substituted by A"  
FT misc\_difference 562  
FT /tag- j  
FT /note- "wild-type A substituted by G"  
FT misc\_difference 583  
FT /tag- k  
FT /note- "wild-type T substituted by C"  
FT misc\_difference 596  
FT /tag- l  
FT /note- "wild-type A substituted by G"  
FT misc\_difference 678  
FT /tag- m  
FT /note- "wild-type A substituted by G"  
FT misc\_difference 681  
FT /tag- n  
FT /note- "wild-type T substituted by C"  
FT misc\_difference 745  
FT /tag- o  
FT /note- "wild-type G substituted by T"  
FT misc\_difference 750  
FT /tag- p  
FT /note- "wild-type G substituted by C"  
FT misc\_difference 765  
FT /tag- q  
FT /note- "wild-type C substituted by T"  
FT misc\_difference 777  
FT /tag- r  
FT /note- "wild-type A optionally substituted by C"  
FT misc\_difference 917  
FT /tag- s  
FT /note- "wild-type T substituted by C"  
FT misc\_difference 1085  
FT /tag- t  
FT /note- "wild-type A substituted by T"  
FT misc\_difference 1256  
FT /tag- u  
FT /note- "wild-type G substituted by A"  
FT misc\_difference 1292  
FT /tag- v  
FT /note- "wild-type T substituted by C"  
FT misc\_difference 1299  
FT /tag- w  
FT /note- "wild-type C substituted by A"  
FT misc\_difference 1325  
FT /tag- x  
FT /note- "wild-type C substituted by A"  
FT misc\_difference 1364  
FT /tag- y  
FT /note- "wild-type C substituted by A"  
FT misc\_difference 1374  
FT /tag- z  
FT /note- "wild-type G substituted by C"  
FT misc\_difference 1392  
FT /tag- aa  
FT /note- "wild-type C substituted by A"  
FT misc\_difference 1454  
FT /tag- ab  
FT /note- "wild-type T substituted by C"  
FT misc\_difference 1478  
FT /tag- ac  
FT /note- "wild-type A substituted by G"  
FT misc\_difference 1617  
FT /tag- ad  
FT /note- "wild-type A substituted by T"  
FT misc\_difference 1744  
FT /tag- ae



FT /\*tag- k  
/note= "wild-type T substituted by C"  
FT misc\_difference 745  
FT /\*tag- l  
/note= "wild-type G substituted by T"  
FT misc\_difference 750  
FT /\*tag- m  
/note= "wild-type G substituted by C"  
FT misc\_difference 765  
FT /\*tag- n  
/note= "wild-type C substituted by T"  
FT misc\_difference 777  
FT /\*tag- o  
/note= "wild-type G substituted by C"  
FT misc\_difference 867  
FT /\*tag- p  
/note= "wild-type A substituted by G"  
FT misc\_difference 917  
FT /\*tag- q  
/note= "wild-type T substituted by C"  
FT misc\_difference 1256  
FT /\*tag- r  
/note= "wild-type G substituted by A"  
FT misc\_difference 1292  
FT /\*tag- s  
/note= "wild-type T substituted by C"  
FT misc\_difference 1325  
FT /\*tag- t  
/note= "wild-type C substituted by A"  
FT misc\_difference 1364  
FT /\*tag- u  
/note= "wild-type C substituted by A"  
FT misc\_difference 1374  
FT /\*tag- v  
/note= "wild-type G substituted by C"  
FT misc\_difference 1392  
FT /\*tag- w  
/note= "wild-type C substituted by A"  
FT misc\_difference 1454  
FT /\*tag- x  
/note= "wild-type T substituted by C"  
FT misc\_difference 1478  
FT /\*tag- y  
/note= "wild-type A substituted by G"  
FT misc\_difference 1854  
FT /\*tag- z  
/note= "wild-type T substituted by C"  
FT misc\_difference 1949  
FT /\*tag- aa  
/note= "wild-type T substituted by A"  
FT misc\_difference 1959  
FT /\*tag- ab  
/note= "wild-type T substituted by C"  
FT misc\_difference 2130  
FT /\*tag- ac  
/note= "wild-type T substituted by C"  
PN W0200029564-A2.  
PD 25-MAY-2000.  
PF 09-NOV-1999; 99MO-US26534.  
PR 12-NOV-1998; 98US-0191468.  
PA (GLIA-) GLIATECH INC.  
PI Gallagher MJ, Burgess LH, Brunden KR;  
XX WPI; 2000-387780/33.  
DR  
XX New nucleic acid encoding a human glycine transporter type 2 (GlyT2),  
PT useful for expressing GlyT2 which can then be used for detecting

PT whether an animal has autoimmune antibodies against a glycine  
PT transporter -  
PS Claim 52; Page -, 175pp; English.  
XX  
XX The present sequence is a mutant version of the coding sequence of the  
CC human glycine transporter type 2 (GlyT2). Glycine is involved in  
CC neurotransmission in the spinal cord and cerebellum, where glycine  
CC receptor inhibition results in pain transmission. The GlyT2 gene and  
CC protein could, therefore, be used to identify inhibitory compounds which  
CC prevent GlyT2 activity, causing a decrease in muscle hyperactivity. They  
CC could be used to treat diseases or conditions associated with muscle  
CC contraction, including muscle spasticity (such as that due to epilepsy,  
CC stroke, head trauma, multiple sclerosis, neuronal cell death, spinal  
CC cord injury and dystonia), myoclonus, epilepsy, Huntington's disease,  
CC amyotrophic lateral sclerosis and pain.  
CC Note: This sequence is not shown in the specification, but is derived  
CC from the hGlyT2 wild type sequence shown in SEQ ID NO: 121 (AA61440).  
CC  
XX  
SQ Sequence 2394 BP; 508 A; 663 C; 628 G; 595 T; 0 other;  
  
Query Match 13.9%; Score 152.6; DB 21; Length 2394;  
Best Local Similarity 55.0%; Pred. No. 6,3e-29;  
Matches 327; Conservative 0; Mismatches 259; Indels 9; Gaps 1;  
  
QY 264 CCTCTGAGAGAACGACACACACAAATTCAGATTATTTCCCAAGTGAAGAA 323  
DB 1742 CCTCCATCGAGACCAATGAGACCTCATCTCAGACAGTTTCCCAAGTGAAGAAC 1801  
QY 324 TGAGGGTCCCATATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 383  
DB 1802 ACAAGCCAGATGTTATCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1861  
QY 384 TGACATGAGCTGGAATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 443  
DB 1862 TCACCTGAGGAGTGAATTTACATGTTTACATGTTTACATGTTTACATGTTTACATG 1921  
QY 444 TTTTAAATTCAGATATACGAGCTAGTGAATTCATGATTTATGAGAGGAGAACAT 503  
DB 1922 TTGTCATCATGTCATTTTGAAGCTGAGAGGAGATCTCATGATGATGCTGCTGCTG 1981  
QY 504 TCATTGAGATACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 563  
DB 1982 TCTGTAGATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2032  
QY 564 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 623  
DB 2033 AAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2092  
QY 624 TGCAATTTGATACCTAATTTATGAGGCAATTCATACCTGATGAGGAGTGTCTTAG 683  
DB 2093 ACCAGTGGAGAACCAATGACCTATGCTCTTACGCTATACCTATGCTGCTGCTGCTG 2152  
QY 684 GCTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743  
DB 2153 GATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2212  
QY 744 TTGAGGCTAAGAGAACATCTTCAACGCTTATAGTTGCTGACAGACGAGCTTAAC 803  
DB 2213 ATCTGCCCCCTGAGAAATTTATGAGAGCTGAAGTTGCTGCTGCTGCTGCTGCTGCT 2272  
QY 804 GGGGTCATACCTGTAACAACATCTGAGGAGAAAGATTAAGACATGATGATGCTGCTG 858  
DB 2273 GGGGCCCATCTTAGCTCAACACCGGGGAGGCTTACAAAGAACATGATGACGCC 2327  
  
RESULT 9  
AAA61439  
ID AAA61439 standard; cDNA; 2394 BP.  
XX  
AC AAA61439;  
XX  
DT 23-FEB-2001 (first entry)

XX DE Human glycine transporter type 2 protein coding sequence #1.  
 XX XX  
 KW Human glycine transporter type 2; GlyT2; muscle spasticity; myoclonus;  
 KW epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;  
 KW dystonia; Huntington's disease; amyotrophic lateral sclerosis; ss.  
 OS Homo sapiens.  
 XX FH  
 FT Key Location/Qualifiers  
 FT CDS 1..2394  
 FT /\*tag= a  
 FT /product= "GlyT2"  
 XX  
 XX MO200029564-A2.  
 XX PD 25-MAY-2000.  
 XX PE 09-NOV-1999; 99MO-US26534.  
 XX PR 12-NOV-1998; 98US-0191468.  
 XX PA (GLIA-) GLIATECH INC.  
 PI Gallagher MJ, Burgess LH, Brunden KR;  
 DR MPI; 2000-387780/33.  
 DR P-PSDB; AAB09896.  
 XX  
 PT New nucleic acid encoding a human glycine transporter type 2 (GlyT2),  
 PT useful for expressing GlyT2 which can then be used for detecting  
 PT whether an animal has autoimmune antibodies against a glycine  
 PT transporter -  
 XX  
 PS Claim 14; Page 157-160; 175pp; English.  
 XX  
 CC The present sequence is one version of the coding sequence of the  
 CC human glycine transporter type 2 (GlyT2). Glycine is involved in  
 CC neurotransmission in the spinal cord and cerebellum, where glycine  
 CC receptor inhibition results in pain transmission. The GlyT2 gene and  
 CC protein could, therefore, be used to identify inhibitory compounds which  
 CC prevent GlyT2 activity, causing a decrease in muscle hyperactivity. They  
 CC could be used to treat diseases or conditions associated with muscle  
 CC contraction, including muscle spasticity (such as that due to epilepsy,  
 CC stroke, head trauma, multiple sclerosis, neuronal cell death, spinal cord  
 CC injury and dystonia), myoclonus, epilepsy, Huntington's disease,  
 CC amyotrophic lateral sclerosis and pain.  
 CC  
 SO Sequence 2394 BP; 505 A; 659 C; 630 G; 600 T; 0 other;  
 Query Match 13.7%; Score 149.4; DB 21; Length 2394;  
 Best Local Similarity 54.6%; Pred. No. 4.2e-28;  
 Matches 325; Conservative 0; Mismatches 261; Indels 9; Gaps 1;  
 QY 264 CCTCTGAAGAAAGCATGACACAAACATTCAGATTATTTCCCAAGATGAGAGAAAA 323  
 DB 1742 CCTCATGACAGACAGACAGCTCCATCTCAGACAGATTCCCAAGTACACGACAC 1801  
 QY 324 TGAAGGTTCCATTAATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 383  
 DB 1802 ACAAGCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1861  
 QY 384 TGACATGAGCTGGAATTTACTGGGTTCACTGATTTGACCACTTCTGCTGATGGGCA 443  
 DB 1862 TGACTCAGGAGTGAATTTACATGTGTTCACCTGTGTGACACATATGCTCCTTATGCC 1921  
 QY 444 TTTTATTTGACCTATACAGACCTATGGAATCTATCTGATTTATGAGGAGACAGAT 503  
 DB 1922 TTGTCTCAATTCGCAATTTTGAAGCTGCTGGGATCTCTTATGTATGCTTGCAGAAAGAT 1981  
 QY 504 TCATTGAGATACAGAAATGATGATGAGCAAAAGAGTGATATCTGCTATGATGATGGA 563  
 DB 1982 TCTGTGAAGATATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2032  
 XX -----TGA 2032

QY 564 GAGCTTCTGCTGTTTGTATTAACGCTATCTCTTTGATTCGAATATTTATCTGCTATGG 623  
 DB 2033 AAGCTGCTGCTGGGATTTGTATTAACCCCAACCTTTATACCTTTATCTGCTGCTTTT 2092  
 QY 624 TGCATTTCTATAGACCTAATTTATGGCCATTTCCATTCCTGAGCTGGAGTTGCTTAG 683  
 DB 2093 ACCAGTGGAAACCCATGACCTATGCTTTACCGTATCTTAACCTGCTCATGCTGCTG 2152  
 QY 684 GCTGTGTATGATTTGTTTCTGATTTATTTGATTCATTTATGCTATGATTAATAA 743  
 DB 2153 GATGGCTAATGCTGCGCTGTTCCGTCATCTGATTCGAATCCCAATTTATGTTGATTAATAATGC 2212  
 QY 744 TTCAGGCTTAAGGAACATCTTTCAACGCTTATTAAGTGTCTGCAACAGCTTCTAAT 803  
 DB 2213 ATCTGGCCCTGGAGATTTATGAGAGGCTGAAGTTGGTGTCTGCTGCAACCGGACT 2272  
 QY 804 GGGGTCATACCTGGAACAAATCGTGGGGAAGATATTAAGCATGATGATGCC 858  
 DB 2273 GGGGCCATTTCTTACTCAACACCGGAGGCTTACAGAAACATGATGATGCC 2327  
 RESULT 10  
 ID AAA61440 standard; cDNA; 2394 BP.  
 XX  
 AC AAA61440;  
 XX  
 DT 23-FEB-2001 (first entry)  
 XX  
 DE Human glycine transporter type 2 protein consensus coding sequence.  
 XX  
 KW Human glycine transporter type 2; GlyT2; muscle spasticity; myoclonus;  
 KW epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;  
 KW dystonia; Huntington's disease; amyotrophic lateral sclerosis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..2394  
 FT /\*tag= a  
 FT /product= "GlyT2"  
 XX  
 XX MO200029564-A2.  
 XX PD 25-MAY-2000.  
 XX PE 09-NOV-1999; 99MO-US26534.  
 XX PR 12-NOV-1998; 98US-0191468.  
 XX PA (GLIA-) GLIATECH INC.  
 PI Gallagher MJ, Burgess LH, Brunden KR;  
 DR MPI; 2000-387780/33.  
 DR P-PSDB; AAB09897.  
 XX  
 PT New nucleic acid encoding a human glycine transporter type 2 (GlyT2),  
 PT useful for expressing GlyT2 which can then be used for detecting  
 PT whether an animal has autoimmune antibodies against a glycine  
 PT transporter -  
 XX  
 PS Claim 13; Page 163-166; 175pp; English.  
 XX  
 CC The present sequence is the consensus coding sequence of the human  
 CC glycine transporter type 2 (GlyT2). The researchers sequenced a number of  
 CC clones from spinal cord and astrocytoma cells to produce this consensus  
 CC version. Glycine is involved in neurotransmission in the spinal cord and  
 CC cerebellum, where glycine receptor inhibition results in pain  
 CC transmission. The GlyT2 gene and protein could, therefore, be used to  
 CC identify inhibitory compounds which prevent GlyT2 activity, causing a  
 CC decrease in muscle hyperactivity. They could be used to treat diseases or

CC conditions associated with muscle contraction, including muscle  
 CC spasticity (such as that due to epilepsy, stroke, head trauma, multiple  
 CC sclerosis, neuronal cell death, spinal cord injury and dystonia),  
 CC myoclonus, epilepsy, Huntington's disease, amyotrophic lateral sclerosis  
 CC and pain.

Sequence 2394 BP; 507 A; 660 C; 628 G; 599 T; 0 other;

Query Match 13.7%; Score 149.4; DB 21; Length 2394;  
 Best Local Similarity 54.6%; Pred. No. 4.2e-28;  
 Matches 325; Conservative 0; Mismatches 261; Indels 9; Gaps 1;

DB 264 CCTGGAAGAAACGATCAACACAAATTCAGATTATTTCCCAAGTGAAGAA 323  
 1742 CCTCATGAGAGACCTAGTACCTTCATCTGACAGAGTTTCCCAAGTGAAGAA 1801  
 324 TGAGGTTCCCAATACCTTTGGGCTGCTGGTTTGTCTTCTTCTGCTGCTG 383  
 1802 ACAAGCCAGTTTACTGCTGGGCTGCTGATTTGTTCTTCATGAGGTTTCCAA 1861  
 384 TGACATGAGGTTGATTTACTGCTGCTGATTTGATGACACCTTCTGATGAGG 443  
 1862 TCACATGAGGTTGATTTACTGCTGCTGATTTGATGACACCTTCTGATGAGG 1921  
 444 TTTTATTCAGCTATCTAGTACGCTAGTATGATGATGATGATGATGATGAT 503  
 1922 TTGTATCATTCATTCATTTTGTAGCTGCTGAGGATCTTATGATGATGATG 1981  
 504 TCATTGAGGATACAGAAATGATGATGATGATGATGATGATGATGATGATG 563  
 1982 TCTGTAAGATATGATGATGATGATGATGATGATGATGATGATGATGATG 2032  
 564 GAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 623  
 2033 AAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2092  
 624 TGCATTTTCATGACCTATTCATTCATTCATTCATTCATTCATTCATTCAT 683  
 2093 ACCAGTGGAGACCCATGACCTATTCATTCATTCATTCATTCATTCATTCAT 2152  
 684 GCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 743  
 2153 GATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2212  
 744 TTCAGGCTTAAGAAACATTCATTCATTCATTCATTCATTCATTCATTCAT 803  
 2213 ATCTGAGCCCTGAGATTATTCATTCATTCATTCATTCATTCATTCATTCAT 2272  
 804 GGGTCCATACCTGCAACATTCGTTGGGAAAGATATTAAGACATGATGATCC 858  
 2273 GGGGCCATTCCTAGCTCAACCGGGGAGCGTTACAAAGACATGATGACCC 2327

RESULT 11  
 AAA91842  
 ID AAA91842 standard; cDNA; 2394 BP.

AC AAA91842;

DT 23-FEB-2001 (first entry)

DE Human glycine transporter type 2 protein mutant coding sequence #2.

KW Human glycine transporter type 2; GlyT2; muscle spasticity; myoclonus;  
 KW epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;  
 KW dystonia; Huntington's disease; amyotrophic lateral sclerosis;  
 KW mutant; mutein; ss.

OS Homo sapiens.

OS Synthetic.

FX Key Location/Qualifiers  
 FT CDS 1..2394

FT /\*tag= a  
 FT /product= "mutant GlyT2 protein"  
 FT misc-difference 70  
 FT /\*tag= b  
 FT /note= "wild-type G substituted by A"  
 FT misc-difference 220  
 FT /\*tag= c  
 FT /note= "wild-type C substituted by T"  
 FT misc-difference 463  
 FT /\*tag= d  
 FT /note= "wild-type A substituted by G"  
 FT misc-difference 562  
 FT /\*tag= e  
 FT /note= "wild-type A substituted by G"  
 FT misc-difference 1085  
 FT /\*tag= f  
 FT /note= "wild-type A substituted by T"  
 FT misc-difference 1292  
 FT /\*tag= g  
 FT /note= "wild-type T substituted by C"  
 FT misc-difference 1299  
 FT /\*tag= h  
 FT /note= "wild-type C substituted by A"  
 FT misc-difference 1617  
 FT /\*tag= i  
 FT /note= "wild-type A substituted by T"  
 FT misc-difference 1744  
 FT /\*tag= j  
 FT /note= "wild-type A substituted by T"  
 FT WO200029564-A2.  
 PD 25-MAY-2000.  
 PF 09-NOV-1999; 99WO-US26534.  
 PR 12-NOV-1998; 98US-0191468.  
 PA (GLIA-) GLIATECH INC.  
 PI Gallagher MJ, Burgess LH, Brunden KR;  
 PI WPI: 2000-387780/33.  
 DR  
 XX  
 PT New nucleic acid encoding a human glycine transporter type 2 (GlyT2),  
 PT useful for expressing GlyT2 which can then be used for detecting  
 PT whether an animal has autoimmune antibodies against a glycine  
 PT transporter -  
 XX  
 PS Claim 12; Page -; 175pp; English.  
 XX  
 CC The present sequence is a mutant version of the coding sequence of the  
 CC human glycine transporter type 2 (GlyT2). Glycine is involved in  
 CC neurotransmission in the spinal cord and cerebellum, where glycine  
 CC receptor inhibition results in pain transmission. The GlyT2 gene and  
 CC protein could, therefore, be used to identify inhibitory compounds which  
 CC prevent GlyT2 activity, causing a decrease in muscle hyperactivity. They  
 CC could be used to treat diseases or conditions associated with muscle  
 CC contraction, including muscle spasticity (such as that due to epilepsy,  
 CC stroke, head trauma, multiple sclerosis, neuronal cell death, spinal  
 CC cord injury and dystonia), myoclonus, epilepsy, Huntington's disease,  
 CC amyotrophic lateral sclerosis and pain.  
 CC Note: This sequence is not shown in the specification, but is derived  
 CC from the hGlyT2 wild type sequence shown in SEQ ID NO: 123 (AAA61441).  
 CC  
 XX  
 SQ Sequence 2394 BP; 506 A; 662 C; 629 G; 597 T; 0 other;

Query Match 13.7%; Score 149.4; DB 21; Length 2394;  
 Best Local Similarity 54.6%; Pred. No. 4.2e-28;  
 Matches 325; Conservative 0; Mismatches 261; Indels 9; Gaps 1;

DB 264 CCTGGAAGAAACGATCAACACAAATTCAGATTATTTCCCAAGTGAAGAA 323  
 1742 CCTCATGAGAGACCTAGTACCTTCATCTGACAGAGTTTCCCAAGTGAAGAA 1801  
 324 TGAGGTTCCCAATACCTTTGGGCTGCTGGTTTGTCTTCTTCTGCTGCTG 383  
 1802 ACAAGCCAGTTTACTGCTGGGCTGCTGATTTGTTCTTCATGAGGTTTCCAA 1861  
 384 TGACATGAGGTTGATTTACTGCTGCTGATTTGATGACACCTTCTGATGAGG 443  
 1862 TCACATGAGGTTGATTTACTGCTGCTGATTTGATGACACCTTCTGATGAGG 1921  
 444 TTTTATTCAGCTATCTAGTACGCTAGTATGATGATGATGATGATGATGAT 503  
 1922 TTGTATCATTCATTCATTTTGTAGCTGCTGAGGATCTTATGATGATGATG 1981  
 504 TCATTGAGGATACAGAAATGATGATGATGATGATGATGATGATGATGATG 563  
 1982 TCTGTAAGATATGATGATGATGATGATGATGATGATGATGATGATGATG 2032  
 564 GAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 623  
 2033 AAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2092  
 624 TGCATTTTCATGACCTATTCATTCATTCATTCATTCATTCATTCATTCAT 683  
 2093 ACCAGTGGAGACCCATGACCTATTCATTCATTCATTCATTCATTCATTCAT 2152  
 684 GCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 743  
 2153 GATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2212  
 744 TTCAGGCTTAAGAAACATTCATTCATTCATTCATTCATTCATTCATTCAT 803  
 2213 ATCTGAGCCCTGAGATTATTCATTCATTCATTCATTCATTCATTCATTCAT 2272  
 804 GGGTCCATACCTGCAACATTCGTTGGGAAAGATATTAAGACATGATGATCC 858  
 2273 GGGGCCATTCCTAGCTCAACCGGGGAGCGTTACAAAGACATGATGACCC 2327

```

Db 1742 CCTCCATCGAGACCATAGTACCTTCATCTCAGACGAGATTCCCAAGTACCTAGCACAC 1801
Qy 324 TGAGGGTTCCCATTAACCTTTGGGCTGCTGCTGTTTCTTCTCTGCTGCTGCTG 383
Db 1802 ACAAGCCAGCTTTTACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1861
Qy 384 TGACTCAGGCTGGAATTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 443
Db 1862 TCACCTCAGGCTGGAATTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1921
Qy 444 TTTTAATTCAGCTATCTGAGCTAGTTGGAATCATCTGATTTATGAGGAGAACAGAT 503
Db 1922 TTGTCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1981
Qy 504 TCATTGAGATACAGAAATGATGATGAGCAAGAGTGGATTTGCTGCTGCTGCTGCTGCTG 563
Db 1982 TCTGTGAAGATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2032
Qy 564 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 623
Db 2032 AAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2092
Qy 624 TGCATTTTCATGAGCTATGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 683
Db 2093 ACCAGTGGAGCCCATGACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2152
Qy 684 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743
Db 2153 GATGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2212
Qy 744 TTCAGGCTAAAGCAACATCTTTCACAGCCTTATGATGCTGCTGCTGCTGCTGCTGCTG 803
Db 2213 ATCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2272
Qy 804 GGGGTCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 858
Db 2273 GGGGCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2327

```

## RESULT 12

AAA92001  
ID AAA92001 standard; cDNA; 2394 BP.

AC AAA92001:

23-FEB-2001 (first entry)

Human glycine transporter type 2 protein mutant coding sequence #5.

Human glycine transporter type 2; GlyT2; muscle spasmodic; myoclonus;  
epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;  
Kw dyslexia; Huntington's disease; amyotrophic lateral sclerosis;  
mutant; mutant; ss.

OS Homo sapiens.

OS Synthetic.

Key Location/Qualifiers

FT CDS

FT /tag- a

FT /product- "mutant GlyT2 protein"

FT /tag- b

FT /note- "wild-type C substituted by T"

FT /tag- c

FT /note- "wild-type A substituted by G"

FT /tag- d

FT /note- "wild-type T substituted by C"

FT /tag- e

FT /note- "wild-type T substituted by C"

PN MO200029564-A2.

XX

```

PD 25-MAY-2000.
XX
XX 09-NOV-1999; 99WO-US26534.
XX
XX 12-NOV-1998; 98US-0191468.
XX
XX (GLIA-) GLIATECH INC.
XX
XX Gallagher MJ, Burgess LH, Brunden KR;
XX WPI: 2000-387780/33.
XX
XX New nucleic acid encoding a human glycine transporter type 2 (GlyT2),
XX useful for expressing GlyT2 which can then be used for detecting
XX whether an animal has autoimmune antibodies against a glycine
XX transporter.
XX
XX Claim 53; Page -; 175pp; English.
XX
XX The present sequence is a mutant version of the coding sequence of the
XX human glycine transporter type 2 (GlyT2). Glycine is involved in
XX neurotransmission in the spinal cord and cerebellum, where glycine
XX receptor inhibition results in pain transmission. The GlyT2 gene and
XX protein could, therefore, be used to identify inhibitory compounds which
XX prevent GlyT2 activity, causing a decrease in muscle hyperactivity. They
XX could be used to treat diseases or conditions associated with muscle
XX contraction, including muscle spasticity (such as that due to epilepsy,
XX stroke, head trauma, multiple sclerosis, neuronal cell death, spinal
XX cord injury and dystonia), myoclonus, epilepsy, Huntington's disease,
XX amyotrophic lateral sclerosis and pain.
XX Note: This sequence is not shown in the specification, but is derived
XX from the hGlyT2 wild type sequence shown in SEQ ID NO: 121 (AAA61440).
XX
SO Sequence 2394 BP; 506 A; 660 C; 629 G; 599 T; 0 other;
Query Match 13.7%; Score 149.4; DB 21; Length 2394;
Best Local Similarity 54.6%; Pred. No. 4.2e-28;
Matches 325; Conservative 0; Mismatches 261; Indels 9; Gaps 1;
Qy 264 CCTGTGAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 323
Db 1742 CCTCCATCGAGACCATAGTACCTTCATCTCAGACGAGATTCCCAAGTACCTAGCACAC 1801
Qy 324 TGAGGGTTCCCATTAACCTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 383
Db 1802 ACAAGCCAGCTTTTACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1861
Qy 384 TGACTCAGGCTGGAATTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 443
Db 1862 TCACCTCAGGCTGGAATTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1921
Qy 444 TTTTAATTCAGCTATCTGAGCTAGTTGGAATCATCTGATTTATGAGGAGAACAGAT 503
Db 1922 TTGTCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1981
Qy 504 TCATTGAGATACAGAAATGATGATGAGCAAGAGTGGATTTGCTGCTGCTGCTGCTGCTG 563
Db 1982 TCTGTGAAGATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2032
Qy 564 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 623
Db 2032 AAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2092
Qy 624 TGCATTTTCATGAGCTATGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 683
Db 2093 ACCAGTGGAGCCCATGACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2152
Qy 684 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743
Db 2153 GATGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2212
Qy 744 TTCAGGCTAAAGCAACATCTTTCACAGCCTTATGATGCTGCTGCTGCTGCTGCTGCTG 803

```





PA (ALIX ) ALLELIX NEUROSCIENCE INC.  
 XX Albert VR, Borden LA, Kowalski LR, McKelvey JF;  
 XX WPI: 1998-169158/15.  
 XX DR P-PSDB: AAM56372.  
 XX  
 PT Human glycine transporter - useful for treating nervous system  
 PT disorders, e.g. pain, myoclonus, etc.  
 XX  
 PS Disclosure: Pages 57-60; 90pp: English.  
 XX  
 CC The present sequence encodes a human GlyT-2 transporter protein. It  
 CC has a different codon usage to the claimed GlyT-2 transporter  
 CC (AAV22903), but encodes the same protein. The cDNA encoding GlyT-2 was  
 CC generated by reverse transcription PCR (RT-PCR) from human spinal cord or  
 CC cerebellum poly A RNA. GlyT-2 is a glycine transporter found  
 CC predominantly in the brain stem and the spinal cord. GlyT-2 regulates the  
 CC levels of glycine, a major neurotransmitter, that preferentially  
 CC influences the activity of N-methyl-D-aspartate (NMDA) receptors and  
 CC strychnine-sensitive glycine receptors. The GlyT-2 protein can be used to  
 CC treat or to diagnose a nervous system disorder or condition, e.g. pain,  
 CC myoclonus, muscle spasm or hyperactivity, epilepsy or spasticity  
 CC associated with stroke, head trauma, neuronal cell death, multiple  
 CC sclerosis, spinal cord injury, dystonia, Huntington's disease or  
 CC amyotrophic lateral sclerosis.  
 CC  
 XX Sequence 2397 BP; 507 A; 669 C; 626 G; 594 T; 1 other;  
 SQ  
 Query Match 13.7%; Score 149.4; DB 19; Length 2397;  
 Best Local Similarity 54.9%; Pred. No. 4.2e-28;  
 Matches 322; Conservative 0; Mismatches 256; Indels 9; Gaps 1;  
 XX  
 QY 272 GAACGATCCACAACAATTCAGATTATTTCCCAAGTGTGAAGAAATGAGGGTT 331  
 DB 1750 GAGACCATATGAGCTCCATCTCAGACAGAGTTTCCCAAGTGTGACACACAGCA 1809  
 QY 332 CCATTAACCTTGGGCTGCTCTGTTGTTTCTCTGCTCTGCTGCTGCTGCTGCTG 391  
 DB 1810 GTCCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1869  
 QY 392 GCTGAATTTTACCGGCTTATCTGATTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTG 451  
 DB 1870 GGTGGAATTTTACCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1929  
 QY 452 GCAGTATCTAGCTGAGCTGATGATCATCTGATTTATGAGAGCAAGATTCATTTGAG 511  
 DB 1930 ATTCGATTTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1989  
 QY 512 GATACGAATATGATTTGAGAGCAAGAGCTGATTTCTGCTGCTGCTGCTGCTGCTG 571  
 DB 1990 GATATGAGATGATGATTTGATTCAGCTTCACTGCTGCTGCTGCTGCTGCTGCTG 2040  
 QY 572 TGGTTTGAATTAACGCTTATCTGATTTGATTCATTTATCTGCTGCTGCTGCTGCTG 631  
 DB 2041 TGGGCTTTGTAACCCCAACATTTTAACTTATCTGCTGCTGCTGCTGCTGCTGCTG 2100  
 QY 632 CATAGACTAATTAATGCGCAATTCATCTGATCTGAGAGGAGTTCCTTATGAGCTGCTG 691  
 DB 2101 GAGCCCATGACCTATGCTCTTACGCTATCTCACTGCTGCTGCTGCTGCTGCTGCTG 2160  
 QY 692 ATGATTTGTTTCTGCTATTTTGGATTCATTTATGCTGCTGCTGCTGCTGCTGCTGCT 751  
 DB 2161 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220  
 QY 752 AAGAGCAATCTTCAACGCTTATTAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 811  
 DB 2221 CCTGGAAGATTTATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280  
 QY 812 TACCTGGAACAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 858  
 DB 2281 TTCTTAGCTCAACACCGGGGAGGCTTACAGAAACATGATGCTGCTGCTGCTGCTGCTG 2327

RESULT 15  
 AAV22903  
 ID AAV22903 standard; cDNA; 840 BP.  
 XX  
 XX AAV22903;  
 AC  
 XX 29-JUL-1998 (first entry)  
 DT  
 XX  
 XX Fragment of the human GlyT-2 gene sequence in clone phg2-7a.  
 DE  
 XX  
 XX Human: GlyT-2 transporter; glycine transporter; neurotransmitter;  
 KW activity: N-methyl-D-aspartate receptor; treatment: diagnosis;  
 KW strychnine-sensitive glycine receptor; myoclonus; muscle spasm; hyperactivity;  
 KW nervous system disorder; head trauma; neuronal cell death;  
 KW multiple sclerosis; spinal cord injury; Huntington's disease;  
 KW amyotrophic lateral sclerosis; ds.  
 KM  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX W09807854-A1.  
 PN  
 XX  
 XX 26-FEB-1998.  
 PD  
 XX  
 XX 20-AUG-1997; 97MO-US14637.  
 PE  
 XX  
 XX 20-AUG-1996; 96US-0700013.  
 PR  
 XX  
 XX (ALIX ) ALLELIX NEUROSCIENCE INC.  
 PA  
 XX Albert VR, Borden LA, Kowalski LR, McKelvey JF;  
 XX WPI: 1998-169158/15.  
 XX DR P-PSDB: AAM56366.  
 XX  
 PT Human glycine transporter - useful for treating nervous system  
 PT disorders, e.g. pain, myoclonus, etc.  
 PS  
 XX Example 1A; Page 43; 90pp: English.  
 XX  
 CC The present sequence encodes a fragment of the human GlyT-2 transporter.  
 CC It corresponds to nucleotides 1540-2379 of the full length sequence  
 CC AAV22903. The cDNA encoding GlyT-2 was generated by reverse  
 CC transcription PCR (RT-PCR) from human spinal cord or cerebellum poly A  
 CC mRNA, using primers AAV22945 and AAV22948. The sequence differs slightly  
 CC from the full length GlyT-2 sequence because samples were collected from  
 CC several donors. GlyT-2 is a glycine transporter found predominantly in  
 CC the brain stem and the spinal cord. GlyT-2 regulates the levels of  
 CC glycine, a major neurotransmitter, that preferentially influences the  
 CC activity of N-methyl-D-aspartate (NMDA) receptors and  
 CC strychnine-sensitive glycine receptors. The GlyT-2 protein can be used to  
 CC treat or to diagnose a nervous system disorder or condition, e.g. pain,  
 CC myoclonus, muscle spasm or hyperactivity, epilepsy or spasticity  
 CC associated with stroke, head trauma, neuronal cell death, multiple  
 CC sclerosis, spinal cord injury, dystonia, Huntington's disease or  
 CC amyotrophic lateral sclerosis.  
 CC  
 XX Sequence 840 BP; 177 A; 219 C; 191 G; 253 T; 0 other;  
 SQ  
 Query Match 13.5%; Score 147.8; DB 19; Length 840;  
 Best Local Similarity 54.7%; Pred. No. 7.5e-28;  
 Matches 321; Conservative 0; Mismatches 257; Indels 9; Gaps 1;  
 XX  
 QY 272 GAACGATCCACAACAATTCAGATTATTTCCCAAGTGTGAAGAAATGAGGGTT 331  
 DB 211 GAGACCATATGAGCTCCATCTCAGACAGAGTTTCCCAAGTGTGACACACAGCA 270  
 QY 332 CCATTAACCTTGGGCTGCTCTGTTGTTTCTCTGCTCTGCTGCTGCTGCTGCTGCTG 391  
 DB 271 GTGTTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 330  
 QY 392 GCTGAATTTTACCGGCTTATCTGATTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTG 451



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 12:07:32 ; Search time 75 Seconds

(Without alignments)  
457.133 Million cell updates/sec

Title: US-09-940-227-83

Perfect score: 1198

Sequence: 1 SEETITTTIDLEPKWKMKM.....VPRKEAGHEIPTVSGSRKPE 216

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneseq/emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneseq/emb1/AA2002.DAT:\*  
24: /SIDSI/gcgdata/geneseq/emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1198	100.0	216	23	ABG60919
2	1176	98.2	414	22	AAW23517
3	1176	98.2	642	21	AAV81943
4	1157	96.6	642	21	AAV81944
5	505.5	42.2	797	19	AAW56381
6	503.5	42.0	797	21	AAW56386
7	503.5	42.0	797	21	AAW56387
8	503.5	42.0	797	21	AAW56388
9	503.5	42.0	797	21	AAW56389

10	503.5	42.0	797	21	AAW56390	Human SC polypept
11	502.5	41.9	280	19	AAW56391	Fragment of the hu
12	502.5	41.9	797	19	AAW56392	Amino acid sequenc
13	502.5	41.9	797	19	AAW56393	Human GLYT-2 trans
14	502.5	41.9	797	19	AAW56394	Human GLYT-2 trans
15	502.5	41.9	797	19	AAW56395	Human GLYT-2 trans
16	502.5	41.9	797	19	AAW56396	Human GLYT-2 trans
17	502.5	41.9	797	19	AAW56397	Human GLYT-2 trans
18	502.5	41.9	797	19	AAW56398	Human GLYT-2 trans
19	502.5	41.9	797	19	AAW56399	Human GLYT-2 trans
20	502.5	41.9	797	19	AAW56400	Human GLYT-2 trans
21	502.5	41.9	797	19	AAW56401	Human GLYT-2 trans
22	502.5	41.9	797	19	AAW56402	Human GLYT-2 trans
23	502.5	41.9	797	19	AAW56403	Human GLYT-2 trans
24	502.5	41.9	797	19	AAW56404	Human GLYT-2 trans
25	502.5	41.9	797	19	AAW56405	Human GLYT-2 trans
26	502.5	41.9	797	19	AAW56406	Human GLYT-2 trans
27	502.5	41.9	797	19	AAW56407	Human GLYT-2 trans
28	502.5	41.9	797	19	AAW56408	Human GLYT-2 trans
29	502.5	41.9	797	19	AAW56409	Human GLYT-2 trans
30	502.5	41.9	797	19	AAW56410	Human GLYT-2 trans
31	501.5	41.9	797	19	AAW56411	Human GLYT-2 trans
32	501.5	41.9	797	19	AAW56412	Human GLYT-2 trans
33	501.5	41.9	797	19	AAW56413	Human GLYT-2 trans
34	501.5	41.9	797	19	AAW56414	Human GLYT-2 trans
35	497.5	41.5	797	19	AAW56415	Human GLYT-2 trans
36	496.5	41.4	797	19	AAW56416	Human GLYT-2 trans
37	491.5	41.0	797	19	AAW56417	Human GLYT-2 trans
38	490.5	40.9	797	19	AAW56418	Human GLYT-2 trans
39	462.5	38.6	194	19	AAW56419	Human GLYT-2 trans
40	447.5	37.4	194	19	AAW56420	Human GLYT-2 trans
41	359	30.0	637	18	AAW11704	Fragment of the hu
42	359	30.0	637	18	AAW11705	Fragment of the hu
43	350.5	29.3	150	21	AAW57224	High affinity Na+-
44	350.5	29.3	150	21	AAW57225	Rat prolilin transp
45	350.5	29.3	150	21	AAW57226	Human glycine tran

#### ALIGNMENTS

RESULT 1  
ID ABG60919 standard; Protein; 216 AA.  
XX ABG60919;  
AC  
XX 13-AUG-2002 (first entry)  
DT  
XX Protein encoded by lung specific gene #9.  
DE  
XX Lung specific gene; gene therapy; vaccine; lung cancer;  
KW cancer staging; cancer monitoring; cancer diagnosis;  
KW imaging lung cancer; metastases.  
XX  
XX Homo sapiens.  
XX WO200218576-A2.  
XX  
XX 07-MAR-2002.  
XX  
XX 27-AUG-2001; 2001WO-US26684.  
XX  
XX 28-AUG-2000; 2000US-228378P.  
XX  
XX (DIAD-) DIADEXUS INC.  
XX  
XX Chen S, Macina RA, Sun Y, Recipon H;  
XX WPI; 2002-434904/46.  
XX  
XX New lung specific genes and their encoded proteins, useful in gene  
PT therapy or as a vaccine for treating lung cancer, as well as for

PT measuring metastases of lung cancer, or staging, monitoring, diagnosing  
 or imaging lung cancer  
 PS Claim 2; Page 205-206; 206pp; English.  
 CC The invention describes a new lung specific gene and its variants. The  
 CC lung specific gene proteins and genes are useful in gene therapy or as a  
 CC vaccine for treating lung cancer. Lung specific genes are also useful for  
 CC staging, monitoring, diagnosing or imaging lung cancer, as well as for  
 CC measuring metastases of lung cancer. This is the amino acid sequence of  
 CC a protein encoded by a lung specific gene described in the invention.  
 SQ Sequence 216 AA;  
 Query Match 100.0%; Score 1198; DB 23; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 2e-127;  
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SEETITTTTIDLPFRVKKMKRVPITLGGCLVFLGLVGYVQAGIYVWHLIDHFCAGWGI 60  
 DB 1 SEETITTTTIDLPFRVKKMKRVPITLGGCLVFLGLVGYVQAGIYVWHLIDHFCAGWGI 60  
 QY 61 LIAAILELVGIIWYIGNGRNFIEDTEMIGAKRWIFLWWRACWFTVITPILLIAIFISLV 120  
 DB 61 LIAAILELVGIIWYIGNGRNFIEDTEMIGAKRWIFLWWRACWFTVITPILLIAIFISLV 120  
 QY 121 QFHRPNYGAIPYPMGVALGCMIVFCIIWIPMAIKIIQAKNIFORLISCCRPASNW 180  
 DB 121 QFHRPNYGAIPYPMGVALGCMIVFCIIWIPMAIKIIQAKNIFORLISCCRPASNW 180  
 QY 181 GPYLEQHRGERYKDMVVPKKEAGHEIPTVSGSRPE 216  
 DB 181 GPYLEQHRGERYKDMVVPKKEAGHEIPTVSGSRPE 216  
 RESULT 2  
 ID AAM23517 standard; Protein; 414 AA.  
 AC AAM23517;  
 DT 12-OCT-2001 (first entry)  
 DE Human EST encoded protein SEQ ID NO: 1042.  
 XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
 KW diagnostics; forensic test; gene mapping; genetic disorder;  
 KW biodiversity; gene therapy; nutrition.  
 OS Homo sapiens.  
 PN WO200154477-A2.  
 PD 02-AUG-2001.  
 PE 25-JAN-2001; 2001WO-US02687.  
 PR 25-JAN-2000; 2000US-0491404.  
 PR 17-JUL-2000; 2000US-0617746.  
 PR 03-AUG-2000; 2000US-0631451.  
 PR 15-SEP-2000; 2000US-0663870.  
 PA (HSE-) HYSEQ INC.  
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
 XX WPI: 2001-476164/51.  
 DR N-PSDB: AAH98176.  
 PT isolated polypeptide for treatment of diseases, diagnostics, raising  
 antibodies and research use

PS Claim 20; Page 796-797; 1275pp; English.  
 CC The present invention provides the protein and coding sequences of novel  
 CC proteins from a variety of organisms, including human, dog, cat, horse,  
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
 CC from the organism of interest. They can be used in diagnostics,  
 CC forensics, gene mapping, identification of mutations, to assess  
 CC biodiversity and for nutritional purposes. The present sequence is a  
 CC protein of the invention.  
 SQ Sequence 414 AA;  
 Query Match 98.2%; Score 1176; DB 22; Length 414;  
 Best Local Similarity 98.6%; Pred. No. 1.4e-124;  
 Matches 213; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 SEETITTTTIDLPFRVKKMKRVPITLGGCLVFLGLVGYVQAGIYVWHLIDHFCAGWGI 60  
 DB 199 SEETITTTTIDLPFRVKKMKRVPITLGGCLVFLGLVGYVQAGIYVWHLIDHFCAGWGI 258  
 QY 61 LIAAILELVGIIWYIGNGRNFIEDTEMIGAKRWIFLWWRACWFTVITPILLIAIFISLV 120  
 DB 259 LIAAILELVGIIWYIGNGRNFIEDTEMIGAKRWIFLWWRACWFTVITPILLIAIFISLV 318  
 QY 121 QFHRPNYGAIPYPMGVALGCMIVFCIIWIPMAIKIIQAKNIFORLISCCRPASNW 180  
 DB 319 QFHRPNYGAIPYPMGVALGCMIVFCIIWIPMAIKIIQAKNIFORLISCCRPASNW 378  
 QY 181 GPYLEQHRGERYKDMVVPKKEAGHEIPTVSGSRPE 216  
 DB 379 GPYLEQHRGERYKDMVVPKKEAGHEIPTVSGSRPE 414  
 RESULT 3  
 ID AAY81943 standard; Protein; 642 AA.  
 AC AAY81943;  
 DT 30-JUN-2000 (first entry)  
 DE Human GUYLIKE protein sequence #1.  
 XX GUYLIKE protein; Glycine transporter protein; human; neuropathic pain;  
 KW chronic pain; postoperative pain; rheumatoid arthritis; neuralgia;  
 KW neuropathy; algesia; nerve injury; ischaemia; neurodegeneration; stroke;  
 KW incontinence; inflammatory disorder; spasticity; myoclonus; epilepsy;  
 KW head trauma; chromosome localisation study; therapy.  
 OS Homo sapiens.  
 PN WO2000014221-A1.  
 PD 16-MAR-2000.  
 PE 03-SEP-1999; 99WO-GB02909.  
 PR 04-SEP-1998; 98GB-0019405.  
 PA (SMIR) SMITHKLINE BEECHAM PLC.  
 PI Duckworth DM, Evans JR;  
 XX WPI: 2000-256973/22.  
 DR N-PSDB: AAA07378.  
 PT New human glycine transporter, known as GUYLIKE, useful in the  
 PT treatment of diseases such as neuropathic pain, chronic pain,  
 PT neuralgia, ischaemia, stroke, incontinence, spasticity and epilepsy,  
 PT comprises a 642 amino acid sequence

PS Claim 2; Page 17-18; 33pp; English.

CC This sequence represents a human glycine transporter (GLYT) protein,  
 CC designated GLYTlike. The GLYTlike polypeptide and the polynucleotide  
 CC encoding it are useful in the treatment of diseases such as neuropathic  
 CC pain, chronic pain, postoperative pain, rheumatoid arthritic pain,  
 CC neuralgia, neuropathies, algesia, nerve injury, ischaemia,  
 CC neurodegeneration, stroke, incontinence, inflammatory disorders,  
 CC spasticity, myoclonus, epilepsy and head trauma. The DNA sequence is  
 CC valuable for chromosome localisation studies. The GLYTlike polypeptide  
 CC and polynucleotide are also useful in diagnostic assays for detecting  
 CC diseases associated with in appropriate GLYTlike activity or levels.

XX Sequence 642 AA:

Query Match 98.2%; Score 1176; DB 21; Length 642;

Best Local Similarity 98.6%; Pred. No. 2.4e-124;

Matches 213; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SEETITTTTODLPFKVKKKRVPTTGGCLVFLGLGVCTGAGIYVWHLIDHFCAGMGI 60  
 DB 427 SIETITTTTODLPFKVKKKRVPTTGGCLVFLGLGVCTGAGIYVWHLIDHFCAGMGI 486  
 QY 61 LIAAILELVGIIWYGGNRFTEDEMMIGAKRMIFWLMWRACWFVTPILLIAIFWSLV 120  
 DB 487 LIAAILELVGIIWYGGNRFTEDEMMIGAKRMIFWLMWRACWFVTPILLIAIFWSLV 546  
 QY 121 QFHRPNYGAIPYPMGVAGLWCMIVFCIWIPIMAIIKIIOAKGNIFORLISCCRPASNW 180  
 DB 547 QFHRPNYGAIPYPMGVAGLWCMIVFCIWIPIMAIIKIIOAKGNIFORLISCCRPASNW 606  
 QY 181 GPYLEQHRGGRYKDMVPPKKEAGHEIPTVSGSRKPE 216  
 DB 607 GPYLEQHRGGRYKDMVPPKKEAGHEIPTVSGSRKPE 642

RESULT 4

AAV81944 standard; Protein: 642 AA.

XX AAV81944;

XX 30-JUN-2000 (first entry)

XX Human GLYTlike protein sequence #2.

XX GLYTlike protein; glycine transporter protein; human; neuropathic pain;  
 KW chronic pain; postoperative pain; rheumatoid arthritic pain; neuralgia;  
 KW neuropathy; algesia; nerve injury; ischaemia; neurodegeneration; stroke;  
 KW incontinence; inflammatory disorder; spasticity; myoclonus; epilepsy;  
 KW head trauma; chromosome localisation study; therapy.

XX Homo sapiens.

XX MO200014221-A1..

XX 16-MAR-2000.

XX 03-SEP-1999; 99MO-GB02909.

XX 04-SEP-1998; 98GB-0019405.

XX (SMIK ) SMITHKLINE BEECHAM PLC.

XX Duckworth DM, Evans JR;

XX WPI; 2000-256973/22.

XX N-PSDB; AAA07379.

XX New human glycine transporter, known as GLYTlike, useful in the  
 PT treatment of diseases such as neuropathic pain, chronic pain,  
 PT neuralgia, ischaemia, stroke, incontinence, spasticity and epilepsy,  
 PT comprises a 642 amino acid sequence -

XX Claim 12; Page 19; 33pp; English.

CC This sequence represents a human glycine transporter (GLYT) protein,  
 CC designated GLYTlike. The GLYTlike polypeptide and the polynucleotide  
 CC encoding it are useful in the treatment of diseases such as neuropathic  
 CC pain, chronic pain, postoperative pain, rheumatoid arthritic pain,  
 CC neuralgia, neuropathies, algesia, nerve injury, ischaemia,  
 CC neurodegeneration, stroke, incontinence, inflammatory disorders,  
 CC spasticity, myoclonus, epilepsy and head trauma. The DNA sequence is  
 CC valuable for chromosome localisation studies. The GLYTlike polypeptide  
 CC and polynucleotide are also useful in diagnostic assays for detecting  
 CC diseases associated with in appropriate GLYTlike activity or levels.

XX Sequence 642 AA:

Query Match 96.6%; Score 1157; DB 21; Length 642;

Best Local Similarity 96.3%; Pred. No. 3.4e-122;

Matches 208; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 SEETITTTTODLPFKVKKKRVPTTGGCLVFLGLGVCTGAGIYVWHLIDHFCAGMGI 60  
 DB 427 SIETITTTTODLPFKVKKKRVPTTGGCLVFLGLGVCTGAGIYVWHLIDHFCAGMGI 486  
 QY 61 LIAAILELVGIIWYGGNRFTEDEMMIGAKRMIFWLMWRACWFVTPILLIAIFWSLV 120  
 DB 487 LIAAILELVGIIWYGGNRFTEDEMMIGAKRMIFWLMWRACWFVTPILLIAIFWSLV 546  
 QY 121 QFHRPNYGAIPYPMGVAGLWCMIVFCIWIPIMAIIKIIOAKGNIFORLISCCRPASNW 180  
 DB 547 QFHRPNYGAIPYPMGVAGLWCMIVFCIWIPIMAIIKIIOAKGNIFORLISCCRPASNW 606  
 QY 181 GPYLEQHRGGRYKDMVPPKKEAGHEIPTVSGSRKPE 216  
 DB 607 GPYLEQHRGGRYKDMVPPKKEAGHEIPTVSGSRKPE 642

RESULT 5

AAW56381 standard; Protein: 797 AA.

XX AAW56381;

XX 29-JUL-1998 (first entry)

XX Human GLYT-2 transporter protein F733S mutant.

XX Human; GLYT-2 transporter; glycine transporter; neurotransmitter;  
 KW activity; N-methyl-D-aspartate receptor; NMDA;  
 KW strychnine-sensitive glycine receptor; treatment; diagnosis;  
 KW nervous system disorder; myoclonus; muscle spasm; hyperactivity;  
 KW epilepsy; spasticity; head trauma; neuronal cell death;  
 KW multiple sclerosis; spinal cord injury; Huntington's disease;  
 KW amyotrophic lateral sclerosis.

XX Synthetic.

XX Homo sapiens.

XX Key Location/Qualifiers

XX MISC-difference 733 /label= F733S

XX W09807854-A1.

XX 26-FEB-1998.

XX 20-AUG-1997; 97WO-US14637.

XX 20-AUG-1996; 96US-0700013.

XX (ALLX ) ALLELIX NEUROSCIENCE INC.

PI	Albert VR,	Borden LA,	Kowalski LR,	McKelvy JF;
XX				
DR	WPL; 1998-169158/15.			
PT	Human glycine transporter - useful for treating nervous system disorders, e.g. pain, myoclonus, etc.			
PS				
XX	Claim 1; Page -: 90pp; English.			
CC	The present sequence represents a human GlyT-2 transporter mutant protein. Glyt-2 is a glycine transporter found predominantly in the brain stem and the spinal cord. GlyT-2 regulates the levels of glycine, a major neurotransmitter, that preferentially influences the activity of N-methyl-D-aspartate (NMDA) receptors and strychnine-sensitive glycine receptors. The GlyT-2 protein can be used to treat or to diagnose a nervous system disorder or condition, e.g. pain, myoclonus, muscle spasm or hyperactivity, epilepsy or spasticity associated with stroke, head trauma, neuronal cell death, multiple sclerosis, spinal cord injury, dystonia, Huntington's disease or amyotrophic lateral sclerosis.			
CC	note: the present sequence does not appear in the specification; it was created using information provided.			
CC				
CC				
SO	Sequence	797 AA:		
OY	Query Match	42.2%; Score 505.5; DB 19; Length 797;		
Dd	Best Local Similarity	46.4%; Pred. No. 3e-48;		
Matches	91; Conservative	34; Mismatches 68; Indels 3; Gaps 1		
OY	3 ETITTTDODLPKKWKRRVPITTGCCCLVFLGLVCVTGAGIYWVHLIDHRCAGMILI 62    :::  :::  :::  :::  :::  :::  :::  :::  :::  :::  :::   584 ETVIVSISDEDPKYLRTKHPPVFTLCCCICFFELMGPMTHGGGIWFOLDVIYAASYALVI 643			
OY	63 AAIETLVGLIIWIYGNGRFLEDTENMIIGAKRWIFLMWRACWEVTPILLIAFIWSLYOF 122    :::  :::  :::  :::  :::  :::  :::  :::  :::  :::  :::   644 IAFELVLGISVYVGIRFCEDIEMMIGQPNLF--MKVCMAFAVPTLTILELCFSFYCW 700			
OY	123 HRPNTGAIPDPDWALGCMIVFCITIWPIMAIKIIOAKNIIFORLICSRPASNMCP 182   :::  :::  :::  :::  :::  :::  :::  :::  :::  :::  :::   701 EPMTGSGRYRYPMSVNLGMILACSIVIWIPIMSIVIKMLAPGRFTERKLVSPOPDMGP 760			
OY	183 VLEOHGREERYKMVP 198 ::      :: : Db	761 FLAQHRGEREKRMNDP 776		
RESULT 6	AAB09896			
ID	AAB09896 standard; Protein; 797 AA.			
XX	AAB09896;			
AC				
DT	23-FEB-2001 (first entry)			
XX				
XX	Human glycine transporter type 2 protein #1.			
KW	Human glycine transporter type 2; GlyT2; muscle spasticity; myoclonus; epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury; dystonia; Huntington's disease; amyotrophic lateral sclerosis.			
OS	Homo sapiens.			
XX				
PX	WO200029564-A2.			
PD	25-MAY-2000.			
PF	09-NOV-1999; 99WO-US26534.			
PR	12-NOV-1998; 98US-0191468.			
PA	(GLIA-) GLIATECH INC.			
Gallaqher MJ,	Burgess LH,	Brunden KR,		

[illegible]

XX WPI: 2000-387780/33.  
 DR N-PSDB: AAA61440.  
 XX  
 PT New nucleic acid encoding a human glycine transporter type 2 (GLYT2),  
 PT useful for expressing GLYT2 which can then be used for detecting  
 PT whether an animal has autoimmune antibodies against a glycine  
 PT transporter -  
 XX  
 PS Claim 10; Page 167-169; 175pp; English.  
 XX  
 CC The present sequence is the consensus protein sequence of the human  
 CC glycine transporter type 2 (GLYT2). The researchers sequenced a number of  
 CC clones of the sequence from spinal cord and astrocytoma cells to produce  
 CC this consensus version. Glycine is involved in neurotransmission in the  
 CC spinal cord and cerebellum, where glycine receptor inhibition results in  
 CC pain transmission. The GLYT2 gene and protein could, therefore, be used  
 CC to identify inhibitory compounds which prevent GLYT2 activity, causing a  
 CC decrease in muscle hyperactivity. They could be used to treat diseases or  
 CC conditions associated with muscle contraction, including muscle  
 CC spasticity (such as that due to epilepsy, stroke, head trauma, multiple  
 CC sclerosis, neuronal cell death, spinal cord injury and dystonia),  
 CC amyoclonus, epilepsy, Huntington's disease, amyotrophic lateral sclerosis  
 CC and pain.  
 CC  
 SO Sequence 797 AA;  
 Query Match 42.0%; Score 503.5; DB 21; Length 797;  
 Best Local Similarity 46.5%; Pred. No. 5e-48;  
 Matches 92; Conservative 33; Mismatches 70; Indels 3; Gaps 1;  
 QY 1 SEETITTTIDLPKVKKKRRVPTTGCCVLFLGLVCTOAGIYVWHLIDHFCAGMI 60  
 Db 582 SIETIYTSIDEPKRYRTKRPVTLGCCCFIFMGPMITGCIYFQVLDYTAASYAL 641  
 QY 61 LIAAILELVGIWYGGNRIEDTEMIGAKRWIFWLMWRCWFVITPILIAIFWISLV 120  
 Db 642 VVIAIFELVGISYVYGLOFCEDEIMMIGOPNIF--WKVCMAFVPTITLFLCSFY 698  
 QY 121 QFHRPNYGAIPYDVGVALGCMIVFCIIPIMAIKITIOAGNIFORLISCRPASNM 180  
 Db 699 QWEPMTYGSYRYPNWSVGLMLACSVIPIPIMFVYIKMLADGRFLERLKLVCSPQPDW 758  
 QY 181 GPYLEQHRGERYKDMVVP 198  
 Db 759 GPFLAQRGERYKRMIDP 776  
 Db  
 RESULT 8  
 AAB23326 standard; Protein: 797 AA.  
 XX  
 AC AAB23326;  
 XX  
 DT 23-FEB-2001 (first entry)  
 XX  
 DE Human glycine transporter type 2 protein mutant #8.  
 XX  
 KW Human glycine transporter type 2; GLYT2; muscle spasticity; myoclonus;  
 KW epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;  
 KW dystonia; Huntington's disease; amyotrophic lateral sclerosis;  
 KW mutant; mutain.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 582 /note="wild-type Thr substituted by Ser"  
 XX  
 PN WO200029564-A2.  
 XX  
 PD 25-MAY-2000.

XX 09-NOV-1999; 99WO-US26534.  
 PF  
 XX 12-NOV-1998; 98US-0191468.  
 PR  
 XX (GLIA-) GLINATECH INC.  
 PA  
 PI Gallagher MJ, Burgess LH, Brunden KR;  
 XX  
 DR WPI: 2000-387780/33.  
 XX  
 CC New nucleic acid encoding a human glycine transporter type 2 (GLYT2),  
 CC useful for expressing GLYT2 which can then be used for detecting  
 CC whether an animal has autoimmune antibodies against a glycine  
 CC transporter -  
 XX  
 PS Claim 7; Page -; 175pp; English.  
 XX  
 CC The present sequence is a mutant version of the protein sequence of the  
 CC human glycine transporter type 2 (GLYT2). Glycine is involved in  
 CC neurotransmission in the spinal cord and cerebellum, where glycine  
 CC receptor inhibition results in pain transmission. The GLYT2 gene and  
 CC protein could, therefore, be used to identify inhibitory compounds which  
 CC prevent GLYT2 activity, causing a decrease in muscle hyperactivity. They  
 CC could be used to treat diseases or conditions associated with muscle  
 CC contraction, including muscle spasticity (such as that due to epilepsy,  
 CC stroke, head trauma, multiple sclerosis, neuronal cell death, spinal  
 CC cord injury and dystonia), myoclonus, epilepsy, Huntington's disease,  
 CC amyotrophic lateral sclerosis and pain.  
 CC Note: This sequence is not shown in the specification, but is derived  
 CC from the hGLYT2 wild type sequence shown in SEQ ID NO: 124 (AAB09898).  
 CC  
 SO Sequence 797 AA;  
 Query Match 42.0%; Score 503.5; DB 21; Length 797;  
 Best Local Similarity 46.5%; Pred. No. 5e-48;  
 Matches 92; Conservative 33; Mismatches 70; Indels 3; Gaps 1;  
 QY 1 SEETITTTIDLPKVKKKRRVPTTGCCVLFLGLVCTOAGIYVWHLIDHFCAGMI 60  
 Db 582 SIETIYTSIDEPKRYRTKRPVTLGCCCFIFMGPMITGCIYFQVLDYTAASYAL 641  
 QY 61 LIAAILELVGIWYGGNRIEDTEMIGAKRWIFWLMWRCWFVITPILIAIFWISLV 120  
 Db 642 VVIAIFELVGISYVYGLOFCEDEIMMIGOPNIF--WKVCMAFVPTITLFLCSFY 698  
 QY 121 QFHRPNYGAIPYDVGVALGCMIVFCIIPIMAIKITIOAGNIFORLISCRPASNM 180  
 Db 699 QWEPMTYGSYRYPNWSVGLMLACSVIPIPIMFVYIKMLADGRFLERLKLVCSPQPDW 758  
 QY 181 GPYLEQHRGERYKDMVVP 198  
 Db 759 GPFLAQRGERYKRMIDP 776  
 Db  
 RESULT 9  
 AAB23327 standard; Protein: 797 AA.  
 XX  
 AC AAB23327;  
 XX  
 DT 23-FEB-2001 (first entry)  
 XX  
 DE Human glycine transporter type 2 protein mutant #9.  
 XX  
 KW Human glycine transporter type 2; GLYT2; muscle spasticity; myoclonus;  
 KW epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;  
 KW dystonia; Huntington's disease; amyotrophic lateral sclerosis;  
 KW mutant; mutain.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX

Key Location/Qualifiers

FT Misc-difference 24 /note- "wild-type Gly substituted by Ser"

FT Misc-difference 74 /note- "wild-type Arg substituted by Trp"

FT Misc-difference 155 /note- "wild-type Ser substituted by Gly"

FT Misc-difference 188 /note- "wild-type Asn substituted by Asp"

FT Misc-difference 362 /note- "wild-type Gln substituted by Leu"

FT Misc-difference 431 /note- "wild-type Val substituted by Ala"

FT Misc-difference 582 /note- "wild-type Thr substituted by Ser"

XX WO200029564-A2.

XX 25-MAY-2000.

XX 09-NOV-1999; 99WO-US26534.

XX 12-NOV-1998; 98US-0191468.

XX (GLIA-) GLIATECH INC.

XX Gallagher MJ, Burgess LH, Brunden KR.

XX WPI; 2000-387780/33.

XX New nucleic acid encoding a human glycine transporter type 2 (GLYT2),

PT useful for expressing GLYT2 which can then be used for detecting

PT whether an animal has autoimmune antibodies against a glycine

PT transporter -

XX Claim 8; Page -: 175bp; English.

XX The present sequence is a mutant version of the protein sequence of the

CC human glycine transporter type 2 (GLYT2). Glycine is involved in

CC neurotransmission in the spinal cord and cerebellum, where glycine

CC receptor inhibition results in pain transmission. The GLYT2 gene and

CC protein could, therefore, be used to identify inhibitory compounds which

CC prevent GLYT2 activity, causing a decrease in muscle hyperactivity. They

CC could be used to treat diseases or conditions associated with muscle

CC contraction, including muscle spasticity (such as that due to epilepsy,

CC stroke, head trauma, multiple sclerosis, neuronal cell death, spinal

CC cord injury and dystonia), myoclonus, epilepsy, Huntington's disease,

CC amyotrophic lateral sclerosis and pain.

CC Note: This sequence is not shown in the specification, but is derived

CC from the hGLYT2 wild type sequence shown in SEQ ID NO: 124 (AAB09898).

XX

SO Sequence 797 AA;

Query Match 42.0%; Score 503.5; DB 21; Length 797;

Best Local Similarity 46.5%; Pred. No. 5e-48;

Matches 92; Conservative 33; Mismatches 70; Indels 3; Gaps 1;

QY 1 SEETITTTODLPPKMKMRVPITLGGCLVFLGLGYCTOAGIYVHLIDHPCAGMGI 60

DB 582 SIEITVTSIDSEPKYLTNRKHPVTLGGCICFFIMGPMITGGIYMQVLDVTAASVAL 641

QY 61 LIAAILELVLGIIVYGNGRFEIDTEMIGAKRWIFWLMWRACWVPTITLAIIFWISLV 120

DB 642 VVIAIEFLVGISVYVIGRCEIDEMIGQPNIF--WKVCMAFVPTITLFLIFCSFSEY 698

QY 121 QEHPRNGAIPYDVGVALGCMIVFCIWIPIMAIKITIOAKGNIFORLISCCRPASNA 180

DB 699 QMEPMYGYSYRYPNMSVGLMLACSVIWIPIMFVIMHLAPGRFIERLKLSCSPQPDW 758

QY 181 GPYLEOHRGERYKDMVVP 198

DB 759 GPFLAQRGRERYKNMIDP 776

RESULT 10

AA82115

ID AAY82115 standard; protein; 797 AA.

XX

XX AAY82115;

AC XX

XX 05-JUN-2000 (first entry)

DT XX

XX Human SC polypeptide glycine transporter SEQ ID NO:2.

DE XX

XX Human; glycine transporter; SC6 polypeptide; screening; diagnosis;

KW identification; neuropathic pain; spasticity; myoclonus; epilepsy;

KW stroke; head trauma; spinal cord injury; dystonia; multiple sclerosis;

KW amyotrophic lateral sclerosis; Huntington's disease;

KW Parkinson's disease.

XX

XX Homo sapiens.

OS XX

XX EP982400-A1.

PN XX

XX 01-MAR-2000.

PD XX

XX 27-OCT-1998; 98EP-0203615.

XX

XX 28-AUG-1998; 98GB-0018890.

PR XX

XX (SMK ) SMITHKLINE BEECHAM PLC.

PA XX

XX Brown AM, Chapman CG, Gloger IS, Evans JR, Cairns W, Herton H;

PI XX

XX WPI; 2000-258613/23.

XX N-PSDB; AA295536.

DR XX

XX New isolated polypeptides and polynucleotides, useful for screening for

PT (ant)agonists and for diagnosing a disease e.g. spasticity, epilepsy,

PT stroke, multiple sclerosis, Huntington's disease or Parkinson's disease

XX

XX Claim 4; Page 17-20; 27bp; English.

XX The present sequence represents the human SC polypeptide glycine

CC transporter. The polypeptide is useful in a method for screening to

CC identify compounds which stimulate or inhibit the function of the

CC polypeptide and in a process for diagnosing a disease or a

CC susceptibility to a disease in a subject related to expression or

CC activity of the polypeptide e.g. treatment of neuropathic pain, other

CC pain, spasticity, myoclonus, epilepsy, stroke, head trauma, spinal cord

CC injury, dystonia, multiple sclerosis, amyotrophic lateral sclerosis,

CC Huntington's disease or Parkinson's disease.

XX

SO Sequence 797 AA;

Query Match 42.0%; Score 503.5; DB 21; Length 797;

Best Local Similarity 46.4%; Pred. No. 5e-48;

Matches 91; Conservative 33; Mismatches 69; Indels 3; Gaps 1;

QY 3 EETITTTODLPPKMKMRVPITLGGCLVFLGLGYCTOAGIYVHLIDHPCAGMGI 62

DB 584 EITVTSIDSEPKYLTNRKHPVTLGGCICFFIMGPMITGGIYMQVLDVTAASVALY 643

QY 63 AAILELVLGIIVYGNGRFEIDTEMIGAKRWIFWLMWRACWVPTITLAIIFWISLV 122

DB 644 IAIIEFLVGISVYVIGRCEIDEMIGQPNIF--WKVCMAFVPTITLFLIFCSFSEY 700

QY 123 HRNNGAIPYDVGVALGCMIVFCIWIPIMAIKITIOAKGNIFORLISCCRPASNA 182

DB 701 EPMYGYSYRYPNMSVGLMLACSVIWIPIMFVIMHLAPGRFIERLKLSCSPQPDW 760

QY 183 YLEOHRGERYKDMVVP 198

DB 761 FLAQRGRERYKNMIDP 776



Dd		188	EMPTGSYNYVPPNNSMWLGNMLACSVINPIFIMVIKMHLPAGFIENKLKVCSPQDPWGP	247
Oy		183	YLEQHRGERYKDMVP	198
		:	:   :	
Dd		248	FLAQRGRERYKNMIDP	263
 RESULT 12				
ID	AAM56372		standard; Protein; 797 AA.	
XX				
AC	AAM56372;			
XX				
DT	29-JUL-1998	(first entry)		
XX				
DE			Amino acid sequence of a human GlyT-2 transporter protein.	
XX				
KW	Human; GlyT-2 transporter; glycine transporter; neurotransmitter;			
KW	activity; N-methyl-D-aspartate receptor; NMDA;			
KW	strychnine-sensitive glycine receptor; treatment; hypnosis;			
KW	nervous system disorder; myoclonus; muscle spasm; hyperactivity;			
KW	epilepsy; spasticity; head trauma; neuronal cell death;			
KW	multiple sclerosis; spinal cord injury; Huntington's disease;			
KW	amyotrophic lateral sclerosis.			
OS	Homo sapiens.			
XX				
PN	WO9807854-A1.			
XX				
PD	26-FEB-1998.			
PF	20-AUG-1997; 97WO-US14637.			
XX				
PR	20-AUG-1996; 96US-0700013.			
XX				
PA	(ALIX ) ALLELIX NEUROSCIENCE INC.			
PI	Albert VR, Borden LA, Kowalski LR, McKelvy JF,			
XX				
DR	WPI; 1998-169158/15.			
XX				
XX	N-PSDB; AAV22309.			
PT			Human glycine transporter - useful for treating nervous system disorders, e.g. pain, myoclonus, etc.	
XX				
PS	Claim 1; Pages 55-56; 90pp; English.			
CC			The present sequence represents a human GlyT-2 transporter protein.	
CC			The cDNA encoding GlyT-2 was generated by reverse transcription PCR (RT-PCR) from human spinal cord or cerebellum polyA mRNA. GlyT-2 is	
CC			a glycine transporter found predominantly in the brain stem and the	
CC			spinal cord. GlyT-2 regulates the levels of glycine, a major	
CC			neurotransmitter, that preferentially influences the activity of	
CC			N-methyl-D-aspartate (NMDA) receptors and strychnine-sensitive glycine	
CC			receptors. The GlyT-2 protein can be used to treat or to diagnose a	
CC			nervous system disorder or condition, e.g. pain, myoclonus, muscle spasm	
CC			or hyperactivity, epilepsy or spasticity associated with stroke, head	
CC			trauma, neuronal cell death, multiple sclerosis, spinal cord injury,	
CC			dystonia, Huntington's disease or amyotrophic lateral sclerosis.	
SQ			Sequence 797 AA;	
 Query Match 41.9%; Score 502.5; DB 19; Length 797; Best Local Similarity 46.4%; Pred. No. 6; Se-48; Matches 91; Conservative 33; Mismatches 69; Indels 3; Gaps 1				
Oy	3	ETTTTIDDLPKVKMRVPITTCGLVLFGLGVCAVQAIGYWHLIDHFCAGGILLI	62	
		:   :         :   :           :   :		
Dd	584	ETVTISIDEFPKYLRTHKPVFLLCCCLCFPMIGCPMTITGGIYMFOLDVTAASVALYI	643	
Oy	63	AAILELVGINIVYGNNRFIEDTEMMIGAKRWITFWLMWRACWFVTPILIAIFINSLYOF	122	
		I I I I I I S V Y Y G I R F C E D I E M M I G P O N I F ---WKYCNAFYVPTITFTLCFSFYOM	700	
Dd	644	IAIEFLVIVSIVYGYRFGCEDEMIIGPOINF---WKYCNAFYVPTITFTLCFSFYOM	700	



CC multiple sclerosis, spinal cord injury, dystonia, Huntington's disease  
CC or amyotrophic lateral sclerosis.  
CC note: the present sequence does not appear in the specification; it was  
CC created using information provided.  
XX

Sequence 797 AA:

Query Match 41.9%; Score 502.5; DB 19; Length 797;  
Best Local Similarity 46.4%; Pred. No. 6.5e-48;

Matches 91; Conservative 33; Mismatches 69; Indels 3; Gaps 1;

QY 3 ETTTITDLPKPKMKRRVPTLGGCLVFLGLVCVTOAGIYVWHLIDHFCAGGII 62

Db 584 ETTTISDDEPKLTKHKVFTLLGCCICFFIMGFPMITOGIYMFQVDTVAASVALVI 643

QY 63 AAIIELVGIITWYGNRFIEDTEMMIGAKRMIFMLMRACWVPTPLIAIFMSLVQF 122

Db 644 IAIIELVGISVYVGLQRFCEDEIMIGFQPNIF---WKYCAFAVPTPLITFLCFSEFYQW 700

QY 123 HRPYGAIPYDPMGVALGCMIVFCITIPMAIITIOAKNIFORLISCCRPASNGP 182

Db 701 EPMITGSRIPNMSVGLMGLACSVIWIPIFVIKMLAPGRFIERLKLVCSPQDMGP 760

QY 183 YLEQHRGERYKDMVVP 198

Db 761 FLAQHRGERYKNMIDP 776

#### RESULT 15

AAAM56375  
ID AAM56375 standard; Protein; 797 AA.

AC AAM56375;

DT 29-JUL-1998 (first entry)

DE Human GlyT-2 transporter protein I279N mutant.

KW Human; GlyT-2 transporter; glycine transporter; neurotransmitter;  
KW activity; N-methyl-D-aspartate receptor; NMDA;  
KW strychnine-sensitive glycine receptor; treatment; diagnosis;  
KW nervous system disorder; myoclonus; muscle spasm; hyperactivity;  
KW epilepsy; spasticity; head trauma; neuronal cell death;  
KW multiple sclerosis; spinal cord injury; Huntington's disease;  
KW amyotrophic lateral sclerosis.

OS Synthetic.

XX Homo sapiens.

XX

XX

Key Location/Qualifiers

FT Misc-difference 279 /label= 1279N

FT /note= "ile replaced with Asn"

PM WO9807854-A1.

XX 26-FEB-1998.

XX 20-AUG-1997; 97MO-US14637.

XX 20-AUG-1996; 96US-0700013.

XX (ALLX ) ALLELIX NEUROSCIENCE INC.

XX Albert VR, Borden LA, Kowalski LR, McKelvey JF;

XX WPI, 1998-169158/15.

XX Human glycine transporter - useful for treating nervous system  
XX disorders, e.g. pain, myoclonus, etc.

XX Claim 1; Page -: 90pp; English.

CC The present sequence represents a human GlyT-2 transporter mutant  
CC protein. GlyT-2 is a glycine transporter found predominantly in  
CC the brain stem and the spinal cord. GlyT-2 regulates the levels of  
CC glycine, a major neurotransmitter, that preferentially influences the  
CC strychnine-sensitive glycine receptors (NMDA) receptors and  
CC used to treat or to diagnose a nervous system disorder or condition,  
CC e.g. pain, myoclonus, muscle spasm or hyperactivity, epilepsy or  
CC spasticity associated with stroke, head trauma, neuronal cell death,  
CC multiple sclerosis, spinal cord injury, dystonia, Huntington's disease  
CC or amyotrophic lateral sclerosis.  
CC note: the present sequence does not appear in the specification; it was  
CC created using information provided.

Sequence 797 AA:

Query Match 41.9%; Score 502.5; DB 19; Length 797;  
Best Local Similarity 46.4%; Pred. No. 6.5e-48;

Matches 91; Conservative 33; Mismatches 69; Indels 3; Gaps 1;

QY 3 ETTTITDLPKPKMKRRVPTLGGCLVFLGLVCVTOAGIYVWHLIDHFCAGGII 62

Db 584 ETTTISDDEPKLTKHKVFTLLGCCICFFIMGFPMITOGIYMFQVDTVAASVALVI 643

QY 63 AAIIELVGIITWYGNRFIEDTEMMIGAKRMIFMLMRACWVPTPLIAIFMSLVQF 122

Db 644 IAIIELVGISVYVGLQRFCEDEIMIGFQPNIF---WKYCAFAVPTPLITFLCFSEFYQW 700

QY 123 HRPYGAIPYDPMGVALGCMIVFCITIPMAIITIOAKNIFORLISCCRPASNGP 182

Db 701 EPMITGSRIPNMSVGLMGLACSVIWIPIFVIKMLAPGRFIERLKLVCSPQDMGP 760

QY 183 YLEQHRGERYKDMVVP 198

Db 761 FLAQHRGERYKNMIDP 776

Search completed: August 27, 2003, 17:34:03  
Job time : 77 secs

**THIS PAGE BLANK (USPTO)**

RNA molecule type: lincRNA

Db 405 ETUVAIVDEGVENETIÖK-KTYVILGVAAGFLLGIPLTSGAGIYVLLMKNYAASFSL 46

Oy 61 LIAALLELVGIIMWYGNRIEDTEMMIGAKRMIFWLMWRACMFVITPILLIAIFWISLV 120  
 Db 464 VVICSCIMCVSIMYIGHRNFQDIOMLGFPPLF---FQICMRFSPTIIFLIFTVI 520  
 Oy 121 QFHRPNYGAIPYDWCVALGCMIVFCIIMPIIMAIKIIOAKN-IFQRLISCCRPASN 179  
 Db 521 QYRPITYNHQYCGMAVAIGFLMASSVICIPLVALFOLCRDGTDLQRLKNAKTRPSRD 580  
 Oy 180 WGPYLEOHNGERYKDMVVPKKEAGHEI 206  
 Db 581 WGPALLEHRTGRVAPTTTSPEDGEFV 607

## RESULT 6

glycine transporter protein - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 20-Aug-1999  
 C:Accession: JH0673; A46204  
 R:Smith, K.E.; Borden, L.A.; Hartig, P.R.; Branchek, T.; Weinschenk, R.L.  
 A:Title: Cloning and expression of a glycine transporter reveal colocalization with NMDA  
 A:Reference number: JH0673; MUID:92265306; PMID:1534013  
 A:Accession: JH0673  
 A:Molecule type: mRNA  
 A:Residues: 1-638 <SMI>  
 A:Cross-references: GB:M88595; NID:g204433; PIDN:AAA41257.1; PID:g204434  
 A:Experimental source: brain  
 R:Gustella, J.; Brecha, N.; Weigmann, C.; Lester, H.A.; Davidson, N.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 7189-7193, 1992  
 A:Title: Cloning, expression, and localization of a rat brain high-affinity glycine tran  
 A:Reference number: A46204; MUID:92357797; PMID:1353889  
 A:Accession: A46204  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 'N', '3', 'GK', '6', 'AKGML', '16-638 <GUA>  
 A:Cross-references: GB:M95413; NID:g204431; PIDN:AAA41256.1; PID:g204432  
 A:Experimental source: C6 glioma cells and tissues  
 A:Note: Sequence extracted from NCBI backbone (NCBI:113358)  
 C:Superfamily: gamma-aminobutyric acid transporter  
 C:Keywords: glycoprotein; transmembrane protein  
 F:41-61/Domain: transmembrane #status predicted <TM1>  
 F:68-88/Domain: transmembrane #status predicted <TM2>  
 F:112-132/Domain: transmembrane #status predicted <TM3>  
 F:222-242/Domain: transmembrane #status predicted <TM4>  
 F:247-267/Domain: transmembrane #status predicted <TM5>  
 F:300-320/Domain: transmembrane #status predicted <TM6>  
 F:329-349/Domain: transmembrane #status predicted <TM7>  
 F:382-402/Domain: transmembrane #status predicted <TM8>  
 F:430-450/Domain: transmembrane #status predicted <TM9>  
 F:466-486/Domain: transmembrane #status predicted <TM10>  
 F:502-522/Domain: transmembrane #status predicted <TM11>  
 F:544-564/Domain: transmembrane #status predicted <TM12>  
 F:169-172,182,188/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 28.8%; Score 344.5; DB 2; Length 638;  
 Best Local Similarity 32.4%; Pred. No. 8.1e-26;

Matches 67; Conservative 54; Mismatches 79; Indels 7; Gaps 4;

Oy 3 ETTITTTIODELPFR--VAKKRVPTLGGCVLFLGLVCTOAGIYVWHLIDHFCAGWGI 60  
 Db 410 ETLVTAIVDEVGEMWIKR-KTYVTGLVAVAGFLGLPLTSQAGITWLLMDYNAASFSL 468  
 Oy 61 LIAALLELVGIIMWYGNRIEDTEMMIGAKRMIFWLMWRACMFVITPILLIAIFWISLV 120  
 Db 466 VVICSCIMCVSIMYIGHRNFQDIOMLGFPPLF---FQICMRFSPTIIFLIFTVI 520  
 Oy 121 QFHRPNYGAIPYDWCVALGCMIVFCIIMPIIMAIKIIOAKN-IFQRLISCCRPASN 179  
 Db 526 QYRPITYNHQYCGMAVAIGFLMASSVICIPLVALFOLCRDGTDLQRLKNAKTRPSRD 585  
 Oy 180 WGPYLEOHNGERYKDMVVPKKEAGHEI 206

Db 586 WGPALLEHRTGRVAPTTTSPEDGEFV 612

## RESULT 7

glycine transporter protein GLYT1a - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Aug-1999  
 C:Accession: S23151; B48716  
 R:Li, Q.R.; Nelson, H.; Mandiyan, S.; Lopez-Corcuera, B.; Nelson, N.  
 FEBS Lett. 305, 110-114, 1992  
 A:Title: Cloning and expression of a glycine transporter from mouse brain.  
 A:Reference number: S23151; MUID:92316254; PMID:1618338  
 A:Accession: S23151  
 A:Molecule type: mRNA  
 A:Residues: 1-633 <LTU>  
 A:Cross-references: EMBL:X67056; NID:951092; PIDN:CAA47440.1; PID:951093  
 R:Li, Q.R.; Lopez-Corcuera, B.; Mandiyan, S.; Nelson, H.; Nelson, N.  
 J. Biol. Chem. 268, 22802-22808, 1993  
 A:Title: Cloning and expression of a spinal cord- and brain-specific glycine transpor  
 A:Reference number: A48716; MUID:94043045; PMID:8226790  
 A:Accession: B48716  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-62 <LI2>  
 A:Genetics:  
 C:Superfamily: gamma-aminobutyric acid transporter  
 C:Keywords: alternative splicing; transmembrane protein

Query Match 28.3%; Score 339.5; DB 2; Length 633;  
 Best Local Similarity 31.9%; Pred. No. 2.5e-25;

Matches 66; Conservative 54; Mismatches 80; Indels 7; Gaps 4;

Oy 3 ETTITTTIODELPFR--VAKKRVPTLGGCVLFLGLVCTOAGIYVWHLIDHFCAGWGI 60  
 Db 405 ETLVTAIVDEVGEMWIKR-KTYVTGLVAVAGFLGLPLTSQAGITWLLMDYNAASFSL 463  
 Oy 61 LIAALLELVGIIMWYGNRIEDTEMMIGAKRMIFWLMWRACMFVITPILLIAIFWISLV 120  
 Db 464 VVICSCIMCVSIMYIGHRNFQDIOMLGFPPLF---FQICMRFSPTIIFLIFTVI 520  
 Oy 121 QFHRPNYGAIPYDWCVALGCMIVFCIIMPIIMAIKIIOAKN-IFQRLISCCRPASN 179  
 Db 521 QYRPITYNHQYCGMAVAIGFLMASSVICIPLVALFOLCRDGTDLQRLKNAKTRPSRD 580  
 Oy 180 WGPYLEOHNGERYKDMVVPKKEAGHEI 206  
 Db 581 WGPALLEHRTGRVAPTTTSPEDGEFV 607

## RESULT 8

dopamine transporter protein - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Aug-1999  
 C:Accession: S20346  
 R:Giros, B.; el Mestikawy, S.; Bertrand, L.; Caron, M.G.  
 FEBS Lett. 295, 149-154, 1991  
 A:Title: Cloning and functional characterization of a cocaine-sensitive dopamine tran  
 A:Reference number: S20346; MUID:92111727; PMID:1765147  
 A:Accession: S20346  
 A:Molecule type: mRNA  
 A:Residues: 1-619 <GIR>  
 A:Cross-references: GB:S76145; NID:g243202; PIDN:AAB21099.1; PID:g243203  
 C:Superfamily: gamma-aminobutyric acid transporter  
 C:Keywords: transmembrane protein

Query Match 27.3%; Score 327; DB 2; Length 619;  
 Best Local Similarity 36.8%; Pred. No. 4.1e-24;  
 Matches 64; Conservative 32; Mismatches 74; Indels 4; Gaps 2;

Oy 3 ETTITTTIODELPFRVAKKRVPTLGGCVLFLGLVCTOAGIYVWHLIDHFCAGWGI 62









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 12:09:04 : Search time 24 Seconds  
(without alignments)  
423.240 Million cell updates/sec

Title: US-09-940-227-83

Perfect score: 1198  
Sequence: 1 SEETITTTTIDLFPRKMKM.....VPRKEAGHEIPTVSGSRKPE 216

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	504.5	42.1	799	1	66A5_RAT
2	502.5	41.9	797	1	66A5_HUMAN
3	361	30.1	636	1	66A7_HUMAN
4	359	30.0	637	1	66A7_RAT
5	345.5	28.8	692	1	66A9_HUMAN
6	344.5	28.8	633	1	66A9_RAT
7	339.5	28.3	633	1	66A9_MOUSE
8	336.5	28.1	638	1	66A9_BOVIN
9	327	27.3	619	1	66A3_MOUSE
10	327	27.3	619	1	66A3_RAT
11	319	26.6	620	1	66A3_HUMAN
12	315	26.3	693	1	66A3_BOVIN
13	308.5	25.8	598	1	66A1_MOUSE
14	308.5	25.8	598	1	66A1_RAT
15	300.5	25.1	598	1	66A1_MOUSE
16	299	25.0	617	1	66A1_HUMAN
17	296	24.7	599	1	66A1_HUMAN
18	293	24.5	617	1	66A2_MOUSE
19	288	24.0	620	1	66A6_BOVIN
20	286.5	23.9	627	1	66A6_RAT
21	286	23.9	615	1	66A2_BOVIN
22	283.5	23.7	602	1	66A2_MOUSE
23	283.5	23.7	602	1	66A2_RAT
24	280	23.4	622	1	66A4_MOUSE
25	278.5	23.2	627	1	66A4_MOUSE
26	277.5	23.2	632	1	66A6_HUMAN
27	273	22.8	614	1	66A6_HUMAN
28	269.5	22.5	641	1	66A6_HUMAN
29	269	22.5	641	1	66A6_HUMAN
30	264	22.0	614	1	66A6_HUMAN
31	260.5	21.7	620	1	66A6_HUMAN
32	259	21.6	614	1	66A6_RAT
33	256.5	21.4	630	1	66A4_BOVIN

34	256.5	21.4	635	1	66A8_BOVIN	018875	bos taurus
35	255.5	21.3	621	1	66A6_RAT	P31643	rattus norv
36	255	21.3	620	1	66A6_CANFA	000589	canis famli
37	254.5	21.2	635	1	66A8_RABIT	P31661	oryctolagus
38	253.5	21.2	621	1	66A6_MOUSE	035316	mus musculu
39	253.5	21.2	621	1	66A6_MOUSE	P31642	mus musculu
40	252.5	21.1	635	1	66A6_MOUSE	P31651	mus musculu
41	252.5	21.1	635	1	66A8_HUMAN	P48029	homo sapien
42	251.5	21.0	635	1	66A8_HUMAN	P28570	rattus norv
43	250.5	20.9	630	1	66A4_MOUSE	060857	mus musculu
44	249.5	20.8	630	1	66A4_RAT	P31652	rattus norv
45	245.5	20.3	569	1	66A6_HUMAN	09nsd5	homo sapien

## ALIGNMENTS

RESULT 1	ID	STANDARD	PRT	799 AA
66A5_RAT	AC	P58295		
16-OCT-2001 (Rel. 40, Created)	DT			
16-OCT-2001 (Rel. 40, Last sequence update)	DT			
15-SEP-2003 (Rel. 42, Last annotation update)	DT			
Sodium- and chloride-dependent glycine transporter 2 (GLYT2) (GLYT-2).	DE			
SIC6A5 OR GLYT2.	GN			
Rattus norvegicus (Rat).	OS			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	OX			
NCBI_Taxid=10116;	RN			
SEQUENCE FROM N.A.	RP			
TISSUE=Brain;	RC			
MEDLINE=94043045; PubMed=8226790;	RX			
Liu Q.-R., Lopez-Corcuera B., Mandiyan S., Nelson H., Nelson N.;	RA			
"Cloning and expression of a spinal cord- and brain-specific glycine	RT			
transporter with novel structural features."	RL			
J. Biol. Chem. 268:22802-22808(1993).				
- FUNCTION: Terminates the action of glycine by its high affinity	CC			
sodium-dependent reuptake into presynaptic terminals. May be	CC			
responsible for the termination of neurotransmission at	CC			
serotonergic-sensitive glycinergic synapses.	CC			
- SUBCELLULAR LOCATION: Integral membrane protein.	CC			
- TISSUE SPECIFICITY: Specifically expressed in spinal cord, brain	CC			
stem, and to a lesser extent in the cerebellum.	CC			
- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)	CC			
family.	CC			
-----				
This SWISS-PROT entry is copyright. It is produced through a collaboration				
between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
the European Bioinformatics Institute. There are no restrictions on its				
use by non-profit institutions as long as its content is in no way				
modified and this statement is not removed. Usage by and for commercial				
entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
-----				
EMBL: L21672; NOT_ANNOTATED_CDS.	DR			
PIR: A48716; A48716.	DR			
InterPro: IPR001175; Na/ntran_symport.	DR			
Pfam: PF00209; SNF: 1.	DR			
PRINTS: PR00176; NANEUSMPORT.	DR			
ProDom: PD000448; Na/ntran_symport; 2.	DR			
PROSITE: PS00610; NA_NEUTROTRAN_SYMPT_1; 1.	DR			
PROSITE: PS00754; NA_NEUTROTRAN_SYMPT_2; 1.	DR			
PROSITE: PS50267; NA_NEUTROTRAN_SYMPT_3; 1.	DR			
Neurotransmitter transporter; Transport; Transmembrane; Glycoprotein;	KW			
Symport.	KW			
DOMAIN	FT			
1	201			
TRANSMEM	202			
230	222			
TRANSMEM	249			
273	293			
TRANSMEM	294			
DOMAIN	395			
TRANSMEM	396			
414				
4				
(POTENTIAL).				
EXTRACELLULAR (POTENTIAL).				
FT				









OY 180 MGPYLEOHRGERKDMVVPKKEACHEI 206  
 DB 581 MGPALLEHRTGKRAFTTPSPEDGEV 607

RESULT 7  
 S6A9 MOUSE STANDARD; PRT: 633 AA.  
 ID S6A9 MOUSE  
 AC P28571:  
 DT 01-DEC-1992 (rel. 24, Created)  
 DT 01-DEC-1992 (rel. 24, Last sequence update)  
 DT 15-SEP-2003 (rel. 42, Last annotation update)  
 DE Sodium- and chloride-dependent glycine transporter 1 (GLYT1) (GLYT-1).  
 GN SLC6A9 OR GLYT1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A. (ISOFORMS GLYT-1A AND GLYT-1B).  
 RX TISSUE=Brain;  
 RX MEDLINE=92316254; PubMed=1618338;  
 RA Liu Q.-R., Nelson H., Mandiyan S., Lopez-Corcuera B., Nelson N.;  
 RT "Cloning and expression of a glycine transporter from mouse brain";  
 RL FEBS Lett. 305:110-114(1992).  
 CC -1- FUNCTION: Terminates the action of glycine by its high affinity  
 CC sodium-dependent reuptake into presynaptic terminals. May play a  
 CC role in regulation of glycine levels in NMDA receptor-mediated  
 CC neurotransmission.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=GLYT-1A;  
 CC IsoId=Q28571-1; Sequence=Displayed;  
 CC Name=GLYT-1B;  
 CC IsoId=P28571-2; Sequence=VSP\_006272;  
 CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)  
 CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X67056; CAA47440.1; -;  
 DR PIR; S23151; S23151.  
 DR MGD; MGI:95760; SLC6A9.  
 DR InterPro: IPR000175; Na/ntran\_symport.  
 DR Pfam; PF00209; SNF\_1  
 DR PRINTS; PR00176; NANEUSMPORT.  
 DR PRODOM; PD000448; Na/ntran\_symport; 2.  
 DR PROSITE; PS00610; NA\_NEUROTRAN\_SYM\_1; 1.  
 DR PROSITE; PS00754; NA\_NEUROTRAN\_SYM\_2; 1.  
 DR PROSITE; PS50267; NA\_NEUROTRAN\_SYM\_3; 1.  
 KW Neurotransmitter transporter; Transport; Transmembrane; Glycoprotein;  
 KW Symport; Amino-acid transport; Alternative splicing;  
 FT DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 36 56 1 (POTENTIAL).  
 FT TRANSSEM 64 83 2 (POTENTIAL).  
 FT TRANSSEM 107 127 3 (POTENTIAL).  
 FT DOMAIN 128 214 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 215 233 4 (POTENTIAL).  
 FT TRANSSEM 242 259 5 (POTENTIAL).  
 FT TRANSSEM 295 312 6 (POTENTIAL).  
 FT TRANSSEM 324 345 7 (POTENTIAL).  
 FT TRANSSEM 378 397 8 (POTENTIAL).  
 FT TRANSSEM 426 444 9 (POTENTIAL).  
 FT TRANSSEM 460 480 10 (POTENTIAL).  
 FT TRANSSEM 501 520 11 (POTENTIAL).

FT TRANSSEM 539 557 12 (POTENTIAL).  
 FT DOMAIN 558 633 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 1 10 MYGKARGML -> MASAGQVATPSPSEQ (in isoform  
 FT GLYT-1B).  
 FT FTrid=VSP\_006272.  
 SQ SEQUENCE 633 AA; 70775 MW; FD64AA561F47093 CRC64;  
 Query Match 28.3%; Score 339.5; DB 1; Length 633;  
 Best Local Similarity 31.9%; Pred. No. 5.1e-23;  
 Matches 66; Conservative 54; Mismatches 80; Indels 7; Gaps 4;

OY 3 ETTITTDLPK--VKKMKVPTTIGCCVLFLGVCYVQAGIYVWHLIDHPCAGMI 60  
 DB 405 ETLVATLDEVGNEMILQK-KTYVTLGVAAVGLGIPLSQAGIYLLMDNAAASFSL 463  
 OY 61 LIAATLELVGIIWYVGNRFEDTEMMIGAKWIFWLMWRACVFPTPIILALFTWSLV 120  
 DB 464 VVISICMVSIMYIGHNRYFDIQMLGFPPPLF--FQICWRFVSPALIFFTIPTVI 520  
 OY 121 QFHRPNCALPPDWGVALGCMIVFCIWPVIAIKTIQAKN-IFORLISCCRPASN 179  
 DB 521 QYRPITYNHQYPPGAVRIGFLMALSSVICIPLALFQLCKTDGDTLLQRLKATKFSRD 580  
 OY 180 MGPYLEOHRGERKDMVVPKKEACHEI 206  
 DB 581 MGPALLEHRTGKRAFTTPSPEDGEV 607

RESULT 8  
 S6A9 BOVIN STANDARD; PRT: 638 AA.  
 ID S6A9 BOVIN  
 AC Q28039; Q28040; Q29420;  
 DT 01-NOV-1997 (rel. 35, Created)  
 DT 01-NOV-1997 (rel. 35, Last sequence update)  
 DT 15-SEP-2003 (rel. 42, Last annotation update)  
 DE Sodium- and chloride-dependent glycine transporter 1 (GLYT1) (GLYT-1).  
 GN SLC6A9.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_Taxid=9913;  
 RN (1)  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
 RA Jones E.M.C.;  
 RU Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: Terminates the action of glycine by its high affinity  
 CC sodium-dependent reuptake into presynaptic terminals. May play a  
 CC role in regulation of glycine levels in NMDA receptor-mediated  
 CC neurotransmission (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=3;  
 CC Name=1; Synonyms=GLYT-1B;  
 CC IsoId=Q28039-1; Sequence=Displayed;  
 CC Name=2; Synonyms=GLYT-1A;  
 CC IsoId=Q28039-2; Sequence=VSP\_006268;  
 CC Name=3; Synonyms=GLYT-1F;  
 CC IsoId=Q28039-3; Sequence=VSP\_006269;  
 CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)  
 CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----





FT TRANSMEM 559 577 12 (POTENTIAL).

FT DOMAIN 578 619 CYTOPLASMIC (POTENTIAL).

FT DISULFID 305 305 INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 181 181 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 188 188 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 196 196 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 204 204 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 204 204 N-LINKED (GLCNAc. . .) (POTENTIAL).

SQ SEQUENCE 619 AA; 68805 MW; F43617BC461E9B CRC64;

Query Match 27.3%; Score 327; DB 1; Length 619;

Best Local Similarity 36.8%; Pred. No. 6.5e-22;

Matches 64; Conservative 32; Mismatches 74; Indels 4; Gaps 2;

QY 3 ETTTITIOQLPFXKKKKRVPITLGGCLVFLGLVCTQAGIYVWHLIDHFCAGWGLI 62

DB 427 ESVITGLVDF-QLLHRRHEFTLGLVATFLSLFCVTNGSIYFTLLDHFAGSTILF 485

QY 63 AALLEVGIIWYGGNREFTEDMMIGAKRMIFWLMRACWEVTPILLIAFIWSLVQF 122

DB 486 GVLIEAIGVAMFYGVQGFSDDIKQMTGQRP--NLYMRCKMLVSPCFLLYVVVSIYTF 542

QY 123 HRPNTGALPYPMGVALGCMIVFCIWIPIWIAIITIKIIOAKNIFQRLISCRP 176

DB 543 RPHVGAIFYPDMANALGWIIATSSMAWPIYATYKFCSLPGSFREKLAYATP 596

RESULT 10

ID S6A3\_RAT STANDARD; PRT; 619 AA.

AC P23977;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Sodium-dependent dopamine transporter (DA transporter) (DAT).

GN SLC6A3.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92054540; PubMed=1948035;

RA Kelly J.E., Loring D., Amara S.G.;

RT Cloning and expression of a cocaine-sensitive rat dopamine transporter.

RT Science 254:578-579(1991).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=92054539; PubMed=1948034;

RA Shimada S., Kitayama S., Lin C.-L., Patel A., Nanthakumar E.,

RA Gregor P., Kuhar M., Uhl G.;

RT Cloning and expression of a cocaine-sensitive dopamine transporter complementary DNA.

RT Science 254:576-578(1991).

RL [3]

RN SEQUENCE FROM N.A.

RX MEDLINE=92111727; PubMed=1765147;

RA Gires B., el Mestikawy S., Bertrand L., Caron M.G.;

RT Cloning and functional characterization of a cocaine-sensitive dopamine transporter.

RT FEBS Lett. 295:149-154(1991).

RN [4]

RP MUTAGENESIS.

RX MEDLINE=92365446; PubMed=1502198;

RA Kitayama S., Shimada S., Xu H., Markham L., Donovan D.M., Uhl G.;

RT Dopamine transporter site-directed mutations differentially alter substrate transport and cocaine binding.

RL Proc. Natl. Acad. Sci. U.S.A. 89:7782-7785(1992).

CC -1- FUNCTION: Amine transporter. Terminates the action of dopamine by its high affinity sodium-dependent reuptake into presynaptic terminals.

CC -1- SUBUNIT: Homooligomer; disulfide-linked. Interacts with PRKCA/B (by similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- MISCELLANEOUS: This protein is the target of psychomotor stimulants such as amphetamines or cocaine.

CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF) family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/or\\_send\\_an\\_email\\_to\\_license@sib-sib.ch](http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch)).

CC EMBL: M80233; AAA41100.1; -

CC EMBL: M80570; AAA73143.1; -

CC EMBL: S76145; AAB21099.1; -

CC PIR: I59558; I59558.

CC PIR: S20346; S20346.

CC InterPro: IPR00175; Na/nttran\_symport.

CC Pfam: PF00209; SNF; 1.

CC PRINTS: PR00176; NANEUSMPORT.

CC ProDom: PD000448; Na/nttran\_symport; 1.

CC PROSITE: PS00610; NA\_NEUOTRAN\_SYM\_1; 1.

CC PROSITE: PS00734; NA\_NEUOTRAN\_SYM\_2; 1.

CC PROSITE: PS50267; NA\_NEUOTRAN\_SYM\_3; 1.

CC Neurotransmitter transporter; Transport; Transmembrane; Glycoprotein; Symport.

KW DOMAIN

FT TRANSMEM 1 68 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 69 89 1 (POTENTIAL).

FT TRANSMEM 96 116 2 (POTENTIAL).

FT TRANSMEM 140 160 3 (POTENTIAL).

FT DOMAIN 161 236 4 (POTENTIAL).

FT TRANSMEM 237 255 5 (POTENTIAL).

FT TRANSMEM 264 281 6 (POTENTIAL).

FT TRANSMEM 317 334 7 (POTENTIAL).

FT TRANSMEM 346 367 8 (POTENTIAL).

FT TRANSMEM 400 419 9 (POTENTIAL).

FT TRANSMEM 446 464 10 (POTENTIAL).

FT TRANSMEM 480 500 11 (POTENTIAL).

FT TRANSMEM 521 540 12 (POTENTIAL).

FT TRANSMEM 559 577 12 (POTENTIAL).

FT TRANSMEM 578 619 12 (POTENTIAL).

FT DISULFID 305 305 INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 181 181 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 188 188 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 196 196 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 204 204 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CONFLICT 597 597 E -> K (IN REF. 3).

SQ SEQUENCE 619 AA; 68746 MW; DF3A30C981095D24 CRC64;

Query Match 27.3%; Score 327; DB 1; Length 619;

Best Local Similarity 36.8%; Pred. No. 6.5e-22;

Matches 64; Conservative 32; Mismatches 74; Indels 4; Gaps 2;

QY 3 ETTTITIOQLPFXKKKKRVPITLGGCLVFLGLVCTQAGIYVWHLIDHFCAGWGLI 62

DB 427 ESVITGLVDF-QLLHRRHEFTLGLVATFLSLFCVTNGSIYFTLLDHFAGSTILF 485

QY 63 AALLEVGIIWYGGNREFTEDMMIGAKRMIFWLMRACWEVTPILLIAFIWSLVQF 122

DB 486 GVLIEAIGVAMFYGVQGFSDDIKQMTGQRP--NLYMRCKMLVSPCFLLYVVVSIYTF 542

QY 123 HRPNTGALPYPMGVALGCMIVFCIWIPIWIAIITIKIIOAKNIFQRLISCRP 176

DB 543 RPHVGAIFYPDMANALGWIIATSSMAWPIYATYKFCSLPGSFREKLAYATP 596

RESULT 11

ID S6A3\_HUMAN STANDARD; PRT; 620 AA.

AC G01959; O14996;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, last sequence update)  
 DT 15-SEP-2003 (Rel. 42, last annotation update)  
 DE Sodium-dependent dopamine transporter (DA transporter) (DAT).  
 GN SLC6A3 OR DAT1.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93061810; PubMed=1359373;  
 RA Vandenbergh D.J., Persico A.M., Uhl G.R.;  
 RT "A human dopamine transporter cDNA predicts reduced glycosylation,  
 RT displays a novel repetitive element and provides racially-dimorphic  
 RT Tag1 RFLPs.";  
 RL Brain Res. Mol. Brain Res. 15:161-166(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Brain;  
 RC MEDLINE=93024326; PubMed=1406597;  
 RA Giros B., el Mestikawy S., Godinot N., Zheng K., Han H.,  
 RA Yang-Feng T., Caron M.G.;  
 RT "Cloning, pharmacological characterization, and chromosome assignment  
 RT of the human dopamine transporter.";  
 RL Mol. Pharmacol. 42:383-390(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Brain;  
 RC MEDLINE=94134051; PubMed=8302271;  
 RA Pristupa Z.B., Wilson J.M., Hoffman B.J., Kish S.J., Niznik H.B.;  
 RT "Pharmacological heterogeneity of the cloned and native human  
 RT dopamine transporter: disassociation of [3H]WIN 35,428 and [3H]GBR  
 RT 12,935 binding.";  
 RL Mol. Pharmacol. 45:125-135(1994).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97444282; PubMed=9300814;  
 RA Kawara T., Kawakami H., Yamamura S.,  
 RT "Structure and organization of the gene encoding human dopamine  
 RT transporter.";  
 RL Gene 195:11-18(1997).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20349832; PubMed=10889531;  
 RA Vandenbergh D.J., Thompson M.D., Cook E.H., Bendahhou E., Nguyen T.,  
 RA Krasowski M.D., Zarrabian D., Comings D., Sellers E.M., Tyndale R.F.,  
 RA George S.R., O'Dowd B.F., Uhl G.R.;  
 RT "Human dopamine transporter gene: coding region conservation among  
 RT normal, Tourette's disorder, alcohol dependence and attention-deficit  
 RT hyperactivity disorder populations.";  
 RL Mol. Psych. 5:283-292(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21200925; PubMed=11304827;  
 RA Greenwood T.A., Alexander M., Keck P.E., McElroy S., Sadovnick A.D.,  
 RA Remick R.A., Kelsoe J.R.;  
 RT "Evidence for linkage disequilibrium between the dopamine transporter  
 RT and bipolar disorder.";  
 RL Am. J. Med. Genet. 105:145-151(2001).  
 RN [7]  
 RP SEQUENCE OF 1-385 FROM N.A.  
 RX MEDLINE=95364623; PubMed=7637582;  
 RA Donovan D.M., Vandenbergh D.J., Perry M.P., Bird G.S., Ingersoll R.,  
 RA Nantakumar E., Uhl G.R.;  
 RT "Human and mouse dopamine transporter genes: conservation of  
 RT 5'-flanking sequence elements and gene structures.";  
 RL Brain Res. Mol. Brain Res. 30:327-335(1995).  
 RN [8]  
 RP SEQUENCE OF 164-255 FROM N.A.  
 RX MEDLINE=92357778; PubMed=135385;  
 RA Bannon M.J., Poosch M.S., Xia Y., Goebel D.J., Cassin B.,  
 RA Kapatos G.;  
 RT "Dopamine transporter mRNA content in human substantia nigra

RT decreases precipitously with age.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:7095-7099(1992).  
 RN [9]  
 RP INTERACTION WITH PRKCABP.  
 RX MEDLINE=21243046; PubMed=11343649;  
 RA Torres G.E., Yao W.-D., Mohn A.R., Quan H., Kim K.-M., Levey A.I.,  
 RA Staudinger J., Caron M.G.;  
 RT "Functional interaction between monoamine plasma membrane transporters  
 RT and the synaptic PDZ domain-containing protein PICK1.";  
 RL Neuron 30:121-134(2001).  
 RN [10]  
 RP INTERCHAIN DISULFIDE BOND.  
 RA Hasting H., Karlin A., Javitch J.A.;  
 RT "Symmetrical homodimer of the human dopamine transporter revealed by  
 RT cross-linking Cys306 at the extracellular end of TM6.";  
 RL Abstr. - Soc. Neurosci. 27:1866-1866(2001).  
 RN [11]  
 RP VARIANT GLN-237.  
 RX MEDLINE=99318093; PubMed=10391209;  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Argille K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.O.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes.";  
 RL Nat. Genet. 22:231-238(1999).  
 RN [12]  
 RP ERRATUM.  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Argille K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.O.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes.";  
 RL Nat. Genet. 22:231-238(1999).  
 CC -1- FUNCTION: Amine transporter. Terminates the action of dopamine by  
 CC its high affinity sodium-dependent reuptake into presynaptic  
 CC terminals.  
 CC -1- SUBUNIT: Homooligomer; disulfide-linked. Interacts with PRKCABP.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- MISCELLANEOUS: This protein is the target of psychomotor  
 CC stimulants such as amphetamines or cocaine.  
 CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)  
 CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; M9670; AAA5770.1; -;  
 DR EMBL; M95167; AAC1720.1; -;  
 DR EMBL; S46955; AAA11754.1; -;  
 DR EMBL; S44626; AAB23443.1; -;  
 DR EMBL; L24178; AAA19560.1; -;  
 DR EMBL; AF321321; AAG33844.1; -;  
 DR EMBL; AF306558; AAG33844.1; JOINED.  
 DR EMBL; AF321320; AAG33844.1; JOINED.  
 DR EMBL; AF306559; AAG33844.1; JOINED.  
 DR EMBL; AF306560; AAG33844.1; JOINED.  
 DR EMBL; AF306561; AAG33844.1; JOINED.  
 DR EMBL; AF306562; AAG33844.1; JOINED.  
 DR EMBL; AF306563; AAG33844.1; JOINED.  
 DR EMBL; AF306564; AAG33844.1; JOINED.  
 DR EMBL; AF119117; AAC50179.2; -;  
 DR EMBL; D88570; BAA22511.1; -;  
 DR EMBL; D88557; BAA22511.1; JOINED.  
 DR EMBL; D88558; BAA22511.1; JOINED.  
 DR EMBL; D88559; BAA22511.1; JOINED.  
 DR EMBL; D88560; BAA22511.1; JOINED.  
 DR EMBL; D88561; BAA22511.1; JOINED.  
 DR EMBL; D88562; BAA22511.1; JOINED.

```

DR EMBL: D88563; BAA22511.1; JOINED.
DR EMBL: D88564; BAA22511.1; JOINED.
DR EMBL: D88565; BAA22511.1; JOINED.
DR EMBL: D88566; BAA22511.1; JOINED.
DR EMBL: D88567; BAA22511.1; JOINED.
DR EMBL: D88568; BAA22511.1; JOINED.
DR EMBL: D88569; BAA22511.1; JOINED.
DR PIR: A48980; A48980.
DR PIR: I57937; I57937.
DR PIR: I84455; I84455.
DR GeneW: HGNC:11049; SLC6A3.
DR MIM: 126455; -.
DR GO: GO:0005737; Cytoplasm; TAS.
DR GO: GO:0005329; Dopamine transporter activity; TAS.
DR GO: GO:0006832; Small molecule transporter; TAS.
DR InterPro: IPR000175; Na/ntan_symport.
DR Pfam: PF00209; SNF; 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR ProDom: PD000448; Na/ntan_symport; 1.
DR PROSITE: PS00610; NA_NEUOTRAN_SYM_1; 1.
DR PROSITE: PS00754; NA_NEUOTRAN_SYM_2; 1.
DR PROSITE: PS0267; NA_NEUOTRAN_SYM_3; 1.
DR Neurotransmitter transporter; Transmembrane; Glycoprotein;
KW Symport; Polymorphism.
FT DOMAIN 1 68
FT TRANSMEM 69 89
FT TRANSMEM 96 116
FT TRANSMEM 140 160
FT DOMAIN 161 237
FT TRANSMEM 238 256
FT TRANSMEM 265 282
FT TRANSMEM 318 335
FT TRANSMEM 347 368
FT TRANSMEM 401 420
FT TRANSMEM 447 465
FT TRANSMEM 481 501
FT TRANSMEM 522 541
FT TRANSMEM 560 578
FT DOMAIN 579 620
FT DISULFID 306 306
FT CARBOHYD 181 181
FT CARBOHYD 188 188
FT CARBOHYD 205 205
FT VARIANT 237 237
FT CONFLICT 35 35
FT CONFLICT 345 345
FT CONFLICT 354 354
FT CONFLICT 367 367
SQ SEQUENCE 620 AA; 68494 MM; CD95009B6BA93108 CRC64;

Query Match 26.6%; Score 319; DB 1; Length 620;
Best Local Similarity 36.2%; Pred. No. 3.4e-21;
Matches 63; Conservative 32; Mismatches 75; Indels 4; Gaps 2;

OY 3 ETTTITDLPKVKMKRVPITLGGCLVFLGLVCVQAGIYWHVLIIDHFCAGWGI 62
DB 428 ESVITGLIDEF-QLLHRRLEFLFLYATFLSLFCVTNGIYVFLDLHFAAGTSLF 486
OY 63 AAILLVGIIITVGNRIETDEMIGAKKIIFMLMRACQFVTPILLIAFIIMSLVOF 122
DB 487 GVLIAIVAMFYGGPSSDIQNTGQRP--SLYRLCKKLVSPELFLFVVVSVITF 543
OY 123 HRPNGAIPYDGMGVALGCMIVFCIIIPIMAIKIIQARGNIFORLISCRP 176
DB 544 RPHHGAIYFPDMALGMLVATSSMAVPIYAAKFKSLPSFSEKLAIVAIAP 597

```

```

DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sodium-dependent dopamine transporter (DA transporter) (DAT).
GN SLC6A3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Eubovae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92107910; PubMed=1722321;
RA Usdin T.B., Mezey E., Chen C., Brownstein M.J., Hoffman B.J.;
RT "Cloning of the cocaine-sensitive bovine dopamine transporter.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:1168-1171(1991).
CC -1- FUNCTION: Amine transporter. Terminates the action of dopamine by
CC its high affinity sodium-dependent reuptake into presynaptic
CC terminals.
CC -1- SUBUNIT: Homooligomer; Disulfide-linked. Interacts with PRKCAP
CC (by similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: This protein is the target of psychomotor
CC stimulants such as amphetamines or cocaine.
CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M80234; -; NOT_ANNOTATED_CDS.
CC PIR: A41617; A41617.
CC InterPro: IPR000175; Na/ntan_symport.
CC Pfam: PF00209; SNF; 1.
CC PRINTS: PR00176; NANEUSMPORT.
CC ProDom: PD000448; Na/ntan_symport; 1.
CC PROSITE: PS00610; NA_NEUOTRAN_SYM_1; 1.
CC PROSITE: PS00754; NA_NEUOTRAN_SYM_2; 1.
CC PROSITE: PS0267; NA_NEUOTRAN_SYM_3; 1.
DR Neurotransmitter transporter; Transmembrane; Glycoprotein;
KW Symport.
FT DOMAIN 1 68
FT TRANSMEM 69 89
FT TRANSMEM 97 116
FT TRANSMEM 140 160
FT TRANSMEM 161 234
FT DOMAIN 235 253
FT TRANSMEM 262 279
FT TRANSMEM 315 332
FT TRANSMEM 344 365
FT TRANSMEM 398 417
FT TRANSMEM 444 462
FT TRANSMEM 478 498
FT TRANSMEM 519 538
FT TRANSMEM 557 575
FT DOMAIN 576 633
FT DISULFID 303 303
FT CARBOHYD 181 181
FT CARBOHYD 196 196
FT CARBOHYD 202 202
SQ SEQUENCE 693 AA; 75691 MM; 002A6D6D0698660 CRC64;

Query Match 26.3%; Score 315; DB 1; Length 693;
Best Local Similarity 31.6%; Pred. No. 8.5e-21;
Matches 71; Conservative 38; Mismatches 98; Indels 18; Gaps 4;

OY 3 ETTTITDLPKVKMKRVPITLGGCLVFLGLVCVQAGIYWHVLIIDHFCAGWGI 62

```

Db 425 ESVTGLADEF-QLLHRRLEFLVLALELLSCFVTCNGIYVFLLDHFAAGSTILE 483  
 QY 63 AAILLVGIITWICGNRFIEDTEMIIGAKRWIFMLWRACWFVITPILLIAIFIMSLVQF 122  
 Db 484 GVLMEVIGVMEFYVWQPSDIDKMTGRP---SLYRLCKEYSPCLLVVVVSYATF 540  
 QY 123 HRPNYGAIPYDVGVALGCMVIFCIIMPIIMAIKIIQAKNIFORLISCCRPASNMG- 181  
 Db 541 RPHYGAIVPEPMATALGMALNASSMSVPIYAAVYKLCISLPSSREKLVAITPETHGR 600  
 QY 182 -----PYLEQ--HREGRYKDMVVPKKEGHEIPYSGSR 213  
 Db 601 VDSCGAPVHAAPLAKGVGRWRKRKSCWVPSRGPGRGPPSPR 645

RESULT 13  
 S6AL\_MOUSE  
 ID S6AL\_MOUSE STANDARD: PRT: 598 AA.  
 AC P31648;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Sodium- and chloride-dependent GABA transporter 1.  
 GN SLC6A1 OR GABT1 OR GAT-1.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP MEDLINE=92333531; PubMed=1631167;  
 RX Lilo O.-R., Mandlyan S., Nelson H., Nelson N.;  
 RL "A family of genes encoding neurotransmitter transporters";  
 CC Proc. Natl. Acad. Sci. U.S.A. 89:6639-6643(1992).  
 CC -1- FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY  
 CC SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: BRAIN.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR  
 CC STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.  
 CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)  
 CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M92377; -; NOT\_ANNOTATED\_CDS.  
 CC EMBL; M92378; -; NOT\_ANNOTATED\_CDS.  
 CC PIR; F46027; F46027.  
 CC MGD; MGI:95627; Gabt1.  
 CC InterPro: IPR000175; Na/ntran\_symport.  
 CC Pfam; PF00209; SNF; 1.  
 CC PRINTS; PR00176; NAMEUSMPORT.  
 CC PRODOM; PD000448; Na/ntran\_symport; 1.  
 CC DR PROSITE; PS00610; NA\_NEUROTRAN\_SYM\_1; 1.  
 CC DR PROSITE; PS00754; NA\_NEUROTRAN\_SYM\_2; 1.  
 CC DR PROSITE; PS50267; NA\_NEUROTRAN\_SYM\_3; 1.  
 CC KW Neurotransmitter transporter; Transport; Transmembrane; Glycoprotein;  
 CC Symport; Multigene family.  
 CC FT DOMAIN 1 52 CYTOPLASMIC (POTENTIAL).  
 CC FT TRANSMEM 53 73 1 (POTENTIAL).  
 CC FT TRANSMEM 81 100 2 (POTENTIAL).  
 CC FT TRANSMEM 124 144 3 (POTENTIAL).  
 CC FT DOMAIN 145 211 EXTRACELLULAR (POTENTIAL).  
 CC FT TRANSMEM 212 229 4 (POTENTIAL).  
 CC FT TRANSMEM 238 255 5 (POTENTIAL).  
 CC FT TRANSMEM 291 308 6 (POTENTIAL).  
 CC FT TRANSMEM 320 341 7 (POTENTIAL).

FT TRANSMEM 374 393 8 (POTENTIAL).  
 FT TRANSMEM 421 439 9 (POTENTIAL).  
 FT TRANSMEM 456 476 10 (POTENTIAL).  
 FT TRANSMEM 497 516 11 (POTENTIAL).  
 FT TRANSMEM 535 553 12 (POTENTIAL).  
 FT DOMAIN 554 598 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 598 AA; 66841 MW; 893C3575DB95BD00 CMC64;

Query Match 25.8%; Score 308.5; DB 1; Length 598;  
 Best local similarity 31.2%; Pred. No. 2,8e-20;  
 Matches 67; Conservative 38; Mismatches 89; Indels 21; Gaps 4;

QY 3 EITITTIQDLPFKMKKRPVITIGCCLVFLGLVCYVQAGIYWHILIDHFCAGCIGL 61  
 Db 401 EGIITLVDEYPLRRNRRELEIAVCIISLITSGIYVFKLFYDSASGMSLL 460  
 QY 62 IAILLVGIITWICGNRFIEDTEMIIGAKRWIFMLWRACWFVITPILLIAIFIMSLVQ 121  
 Db 461 FLVFEVYSISMEFYGNRFYDNIQEMGSRCT---WKLCMSFTPIYAGVFLFAVO 517  
 QY 122 HRPNYGAIPYDVGVALGCMVIFCIIMPIIMAIKIIQAKNIFORLISCCRPASNMG 181  
 Db 518 MTPLYGSIYVFPKMGQGVGMALMSVLIPIGYAAVFLTKSLKORLQVMIQPS---- 573  
 QY 182 PYLEQHGERYKDMVVPKKEGHEIPYSGSRKE 216  
 Db 574 -----EDIVRP--ENGPEOPAGSSAKE 595

RESULT 14  
 S6AL\_MOUSE  
 ID S6AL\_MOUSE STANDARD: PRT: 599 AA.  
 AC P23978;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Sodium- and chloride-dependent GABA transporter 1.  
 GN SLC6A1 OR GABT1 OR GAT-1.  
 OS Rattus norvegicus (Rat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP MEDLINE=90378307; PubMed=1975955;  
 RX MEDLINE=90378307; PubMed=1975955;  
 CC TISSUE-BRAIN:  
 CC Gustella J., Nelson N., Nelson H., Czyzyk L., Keynan S.,  
 RA Medel M.C., Davidson N., Lester H.A., Kanner B.I.;  
 RT "Cloning and expression of a rat brain GABA transporter";  
 RL Science 249:1303-1306(1990).  
 CC [2]  
 CC TOPOLOGY:  
 CC MEDLINE=97313439; PubMed=9169433;  
 CC Clark J.A.;  
 RT "Analysis of the transmembrane topology and membrane assembly of the  
 RT GAT-1 gamma-aminobutyric acid transporter";  
 RL J. Biol. Chem. 272:14695-14704(1997).  
 CC -1- FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY  
 CC SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR  
 CC STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.  
 CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)  
 CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

[illegible]

Query Match	Best Local	Similarity	Score	DB 1	Length	598
Matches	66	Conservative	38	Mismatches	90	Indels
QY	3	ETITTTDDLPFKWKKRRVITITGCLVLELLGLVCTYQAGIYVHLLDHFCA-GMCI	1	1	1	1
DB	401	EGFTTALVDEYSRLLRNRRELFIAVACVTSVLSIGNTGGIYVFKLEDFYSSASGSL	1	1	1	1
QY	62	IAAILELGLIWTGNGRFIDETEMGICAKWFIWLMARACWFYITPILLAIFRWISVQ	1	1	1	1
DB	461	FLVPEECYSISWFGVNRFDYNIQMGSSRCI---WKKLMSFTPTIIVAGVFLFSVQ	1	1	1	1

Wed Aug 27 18:31:21 2003

us-09-940-227-83.rsp

Page 13

[illegible]

Search completed: August 27, 2003, 17:34:41  
Job time : 25 secs

**THIS PAGE BLANK (USPTO)**

---



QY	264	CCCTCGAAGAAGCATCACAAACAATTCAGATTATTTATCCCAAGATGATGAAGAAA	323
Db	1742	CCCTCATGAGACCATATGTCACCTCCATCTCAGACGAATTTCCCAAGATCACTGACACAC	1801
QY	324	TGAGGGTTCGCATAACTTTGGGGTCGTCCTGGTTTCTTCTCCTTGGCTCGCTGTGG	383
Db	1802	ACAAGCCAGTGTTCCTCTCTGGGGCTGCATATTTGTTCTTCATCATGGGGTTTCCAAATGA	1863
QY	384	TGACTCAGGCTGGAATTTACTGGGGTTCATCTGATTTGACCACTTCGTCTGGATGGGCA	443
Db	1862	TCACTCAGGGTGGAAATTTACATGTTTACGCTTGTGGACACCTATGCTCCTCCTCATGGCC	1922
QY	444	TTTTAAATTCGACCTATCTGAGCTAGTTGGAAATCATCTGCATTTATGAGGGAACAGAT	503
Db	1922	TTGTCATCATTTGCCATTTTGGAGTCGGGGGAACTCTTATGTATGCTGTCGCAAAAT	1981
QY	504	TCATTTGAGATACGAATGATGATTTGGAGCAAGAGGTGATATTTGGCTATGATGTGA	563
Db	1982	TCGTGTGAAGATATATGAGATGATGATTTGGATTCCAGCCCAATCATCTTC-----TGGCA	2033
QY	564	GAGCTTCTGCTGTTTGTATTAACGCTTATCCCTTTTGATTTGCAATATTTATCTGGTCATTTGG	623

DB 2033 AAGTCGTGGGCAATTTTAAACCCACACATTTTAACCTTTACCTTGGCTTACGCTTT 2092  
OY 624 TGCATTTTCATAGACCTTAATTAATGAGCGCAATTCATACCGTGGGAGAGTTGCTTAA 683  
DB 2093 AACATGAGGAACCCATGACATGACATGCTCTTACCGCTATCCCTAATGATGCTGCTG 2152  
OY 684 GCTGTGATGATGTTTCTTCTGATATTTTGGATTTTGCATTTTATGCTATATGCTATTA 743  
DB 2153 GATGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2212  
OY 744 TTCAGGCTAAAGAAACATCTTTCAACGCTTATAGTTGCTGACAGACGCTTCTACT 803  
DB 2213 ATCTGGCCCCCGAAGATTTTATGAGAGCTGCAATTTGTTGCTGCGCACAGCGGACT 2272  
OY 804 GGGGTCCATACCTGGAACACATCTGTTGAGGAAAGATTAATTAAGACATGTTGCTCC 858  
DB 2273 GGGGCCCATTTCTTACCTCAACACCGCGGAGCGCTTACAGAAACATGATCGACCC 2327

## RESULT 2

US-09-191-468-121  
; Sequence 121, Application US/09191468A  
; Patent No. 6416975  
; GENERAL INFORMATION:  
; APPLICANT: Gallagher, Michael J.  
; APPLICANT: Burgess, Loyd R.  
; APPLICANT: Brunden, Kurt R.  
; TITLE OF INVENTION: Human Glycine Transporter Type 2  
; FILE REFERENCE: 12311US01  
; CURRENT APPLICATION NUMBER: US/09/191,468A  
; CURRENT FILING DATE: 1998-11-12  
; NUMBER OF SEQ ID NOS: 124  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 121  
; LENGTH: 2394  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2391)  
; OTHER INFORMATION: Seq50 Consensus; nt 1-2394  
US-09-191-468-121

Query Match 13.7%; Score 149.4; DB 4; Length 2394;  
Best Local Similarity 54.6%; Pred. No. 2,1e-30;  
Matches 325; Conservative 0; Mismatches 261; Indels 9; Gaps 1;

OY 264 CCTCTGAGAAACGATCACACAAATTCAGATTATTTCCCAAGATGAGAGAAA 323  
DB 1742 CCTCATGAGAACATGACATGACCTCATCTCAGACAGATTCCCAAGTACCTAGCAGAC 1801  
OY 324 TGAGGCTTCCATTAATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 383  
DB 1802 ACAAGCAGTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1861  
OY 384 TGACGAGCTGGAATTTACTGGGTCATCTGATGACACATCTGCTGATGAGGGA 443  
DB 1862 TCACTGAGGGAATTTACATGTTTCACTTGTGACACCTATGCTGCTCTTATGCC 1921  
OY 444 TTTTATTCAGCTATCTAGCTGAGCTGATGATCTGATCTGATTTATGAGGGAACAGAT 503  
DB 1922 TTTTATTCAGCTATCTAGCTGAGCTGATGATCTGATCTGATTTATGAGGGAACAGAT 1981  
OY 504 TCATTTGAGATACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 563  
DB 1982 TCTGTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2032  
OY 564 GAGCTGTGCTGTTGATTAATGAGCTGATCTGATGATGATGATGATGATGATGATGAT 623  
DB 2033 AAGTCGTGGGCAATTTTAAACCCACACATTTTAAACCTTTTACCTTGGCTTACGCTTT 2092  
OY 624 TGCATTTTCATAGACCTTAATTAATGAGCGCAATTCATACCTGAGCTGAGGAGTTGCTT 683

DB 2093 AACATGAGGAACCCATGACATGACATGCTCTTACCGCTATCCCTAATGATGCTAGGNGCTG 2152  
OY 684 GCTGTGATGATGTTTCTTCTGATATTTTGGATTTTCAATTAATGCTATCAATTA 743  
DB 2153 GATGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2212  
OY 744 TTCAGGCTAAAGAAACATCTTTCAACGCTTATAGTTGCTGACAGACGCTTCTACT 803  
DB 2213 ATCTGGCCCCCGAAGATTTTATGAGAGCTGCAATTTGTTGCTGCGCACAGCGGACT 2272  
OY 804 GGGGTCCATACCTGGAACACATCTGTTGAGGAAAGATTAATTAAGACATGTTGCTCC 858  
DB 2273 GGGGCCCATTTCTTACCTCAACACCGCGGAGCGCTTACAGAAACATGATCGACCC 2327

## RESULT 3

US-08-700-013B-14  
; Sequence 14, Application US/08700013B  
; Patent No. 5919653  
; GENERAL INFORMATION:  
; APPLICANT: Albert, Vivian R.  
; APPLICANT: Kowalski, Leslie R.Z.  
; APPLICANT: Borden, Laurence A.  
; APPLICANT: McKeivly, Jeffrey F.  
; TITLE OF INVENTION: Human Glycine Transporter  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert Price & Rhoads  
; STREET: 997 Lenox Drive, Building 3, Suite 210  
; CITY: Lawrenceville  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08543  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/700,013B  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bloom, Allen  
; REGISTRATION NUMBER: 29,135  
; REFERENCE/DOCKET NUMBER: 317743-108  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 609-520-3214  
; TELEFAX: 609-520-3259  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 840 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-700-013B-14

Query Match 13.5%; Score 147.8; DB 2; Length 840;  
Best Local Similarity 54.7%; Pred. No. 3.9e-30;  
Matches 321; Conservative 0; Mismatches 257; Indels 9; Gaps 1;

OY 272 GAAGAGTACAAACAATTAAGATTTATTTCCCAAGATGATGAAGAAATGAGGTT 331  
DB 211 GAGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 270  
OY 332 CCCATTAATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391  
DB 271 GTGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 330

```

Oy 392 GCTGGAATTTACTGGGTTTCATCTGATTCAGCCATCTCTGCTGATGGGCAATTTAAT 451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 GGTGAATTTAGCATGTTTACGCTTGTGACACATTCGTGCTCTATGCCCTGTCTATC 390
Oy 452 GCAGCTTACTGAGCATGTTGGAATTCATCTGATTTATGAGGGAACATTCATGAG 511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 ATTCGCAATTTTGAAGCTGCTGGGATCTTATGATGATGCTGCAAAAGATTCGTGA 450
Oy 512 GATACAGAAATGATGATGAGCAAAAGAGTGATATTCCTGCTATGCTGAGAGCTTGC 571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 GATATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 501
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 572 TGGTTTGAATTAAGCCATCTCTTTCATTCATTAATTTATCTGCTATGCTGCAATTT 631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 TGGGCAATTTTGAACCCCAACCATTTTATCCTTTCCTTTCAGCTTTTACAGTGG 561
Oy 632 CATGACCTAATTTATGGCGCAATTCATACCTGACGAGGAGTTGCTTTAGGCTGTGT 691
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 562 GAGCCATGACCTTATGCTTACCGCTATCTCACTGATGCTGATGCTGATGCTG 621
Oy 692 ATGATTTGTTCTGCAATTTATTTGATTCATTTATGCTATCAATTAATTAATTCAGCT 751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 622 ATGCTCCGCTTTCGCTCATCTGATCCCAATTTATGTTGATTAATTAATTCATCTG 681
Oy 752 AAAGAAACATCTTTCAACGCTTATTAAGTCTGACAGACCGCTTCACTGAGGCTCA 811
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 682 CCTGGAAGATTTATGAGAGCTGAAGTTGCTGCTGCTGCTGCTGCTGCTGCTG 741
Oy 812 TACCTGGAACAACTCTGGGGAAGATTAATTAACATGCTGCTGCTGCTGCTGCTG 858
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 742 TTCCTGCTCAACACCGCGGGAGCGTTACAAAGAAATGATGAGCC 788

```

## RESULT 4

```

US-08-700-013B-16
: Sequence 16 Application US/08700013B
: Patent No. 5919653
: GENERAL INFORMATION:
: APPLICANT: Albert, Vivian R.
: APPLICANT: Kowalski, Leslie R.Z.
: APPLICANT: Borden, Laurence A.
: APPLICANT: McKelvy, Jeffrey F.
: TITLE OF INVENTION: Human Glycine Transporter
: NUMBER OF SEQUENCES: 41
: CORRESPONDENCE ADDRESS:
: STREET: 997 Lenox Drive, Building 3, Suite 210
: CITY: Lawrenceville
: STATE: NJ
: COUNTRY: USA
: ZIP: 08543
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/700,013B
: FILING DATE:
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Bloom, Allen
: REGISTRATION NUMBER: 29,135
: REFERENCE/DOCKET NUMBER: 317743-108
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 609-520-3214
: TELEFAX: 609-520-3259
: TELEX:
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:

```

```

: LENGTH: 840 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
US-08-700-013B-16

```

```

Query Match 13.5%; Score 147.8; DB 2; Length 840;
Best Local Similarity 54.7%; Pred. No. 3.9e-30;
Matches 321; Conservative 0; Mismatches 257; Indels 9; Gaps 1;

```

```

Oy 272 GAACGATCACACAAATTCATCAAGATTTATTCGCCAAGATGATGAAGAAATGAGGTT 331
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 GAGACCAATTAAGACCTTCATCTCAGAGAGTTTCCCAAGTACACGACACACAGCA 270
Oy 332 CCATTAATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 GTGTTACTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 330
Oy 392 GCTGGAATTTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 GGTGAATTTTACATGTTTACGCTTGTGACACCTATGCTGCTGCTGCTGCTGCTGCTG 390
Oy 452 GCAGCTTACTGAGCATGTTGGAATTCATCTGATTTATGAGGGAACAGATTCATGAG 511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 ATTCGCAATTTTGAAGCTGCTGGGATCTTATGATGATGCTGCAAAAGATTCGTGA 450
Oy 512 GATACAGAAATGATGATGAGCAAAAGAGTGATATTCCTGCTATGCTGAGAGCTTGC 571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 GATATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 501
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 572 TGGTTTGAATTAAGCCATCTCTTTCATTCATTAATTTATCTGCTATGCTGCAATTT 631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 TGGGCAATTTTGAACCCCAACCATTTTATCCTTTCCTTTCAGCTTTTACAGTGG 561
Oy 632 CATGACCTAATTTATGGCGCAATTCATACCTGACGAGGAGTTGCTTTAGGCTGTGT 691
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 562 GAGCCATGACCTTATGCTTACCGCTATCTCACTGATGCTGATGCTGATGCTG 621
Oy 692 ATGATTTGTTCTGCAATTTATTTGATTCATTTATGCTATCAATTAATTAATTCAGCT 751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 622 ATGCTCCGCTTTCGCTCATCTGATCCCAATTTATGTTGATTAATTAATTCATCTG 681
Oy 752 AAAGAAACATCTTTCAACGCTTATTAAGTCTGACAGACCGCTTCACTGAGGCTCA 811
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 682 CCTGGAAGATTTATGAGAGCTGAAGTTGCTGCTGCTGCTGCTGCTGCTGCTG 741
Oy 812 TACCTGGAACAACTCTGGGGAAGATTAATTAACATGCTGCTGCTGCTGCTGCTG 858
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 742 TTCCTGCTCAACACCGCGGGAGCGTTACAAAGAAATGATGAGCC 788

```

## RESULT 5

```

US-09-191-468-123
: Sequence 123 Application US/09191468A
: Patent No. 6416975
: GENERAL INFORMATION:
: APPLICANT: Gallagher, Michael J.
: APPLICANT: Burgess, Loyd R.
: APPLICANT: Brunden, Kurt R.
: TITLE OF INVENTION: Human Glycine Transporter Type 2
: FILE REFERENCE: 1231US01
: CURRENT APPLICATION NUMBER: US/09/191,468A
: CURRENT FILING DATE: 1998-11-12
: NUMBER OF SEQ ID NOS: 124
: SOFTWARE: Patentl Ver. 2.0
: SEQ ID NO 123
: LENGTH: 2394
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(2391)
: OTHER INFORMATION: SEQ ID NO:26 [W098/07854 (PCT/US97/14637)] All611x

```



[illegible]

	Query Match	13.5%	Score 147.8	DB 3	Length 2863
	Best Local Similarity	54.7%	Pred. No. 6	Le-30	
	Matches 321	Conservative 0	Mismatches 257	Indels 9	Gaps 1
QY	272	GAAACGATCAACACAAATTCAGATTATTTCCCAAAGTAGTAMGMAAATGAGGGTT	331		
Db	2005	GAGCCCAATAGTAGACCTCCATCTCAGACGAGATTGCCAAGTAGCTACGACACACAAGCCA	206		
QY	332	CCCATACCTTGGGCGCTGCTGGTGGTTTGGTTTCCCTGGCTCAGCTGTGTGACTAG	391		
Db	2065	GTGTTTACTCTGGGCTCTGTCATTTGTTTCTTCATCATGGGTTTTCCANTGATCACTAG	2124		
QY	392	GCTGGAAATTACTGGGGTTTCATCTGATTGACCACTTCGTGCTGATGGGCAATTTAAT	451		
Db	2125	GGTGGAAATTACATGTTTCAGCTGTGGACACCTATGCTGCTCTATAGCCCTGTGATC	2184		
QY	452	GCAGCTTACTGGGCTAGTTGGAAATCATCTGATTTATGAGGGAACAGATTCAATGAG	511		
Db	2185	ATTGCCATTTTGGAGCTCGTGGGATCTCTTATGTATGAGCTTGCACAAAGATTTGTGAA	2244		
QY	512	GATACAGAAATGATGATTGGAGCAAAAGAGGTGATATCTGGCATATGGTGAGAGCTTGC	571		
Db	2245	GATATAGAGATGATGATTGGATTCCACCCCTAACAATCTC-----TGGAAAGTCTGC	2295		
QY	572	TGGTTTGTATTTACGCTTATCCTTTTGATTTGCAATATTTATCTGTGTAATGGTGCATTT	631		
Db	2296	TGGGCAATTTGTAAACCCCAACCAATTTTAACTTTATCCTTGGCTTCAGCTTTTAACTAGTG	2355		
QY	632	CATAGACCTTAATATGAGGCGCAATTCCTACCTGACGGGAGAGTGGCTTAAAGCTGGTGT	691		
Db	2356	GAGCCCATGACATATGAGCTCTTACCGCTATACCTAATCAGTGCATGAGTGTCTCGATGGCTA	2415		
QY	692	ATGATTTGTTTCTGCATTAATTTGGATTCCAAATTATGGCTATCAATAAATAATTCAGGCT	751		





## TELECOMMUNICATION INFORMATION:

TELEPHONE: 609-520-3214  
TELEFAX: 609-520-3259

TELEX:

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 2403 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-700-013B-26

Query Match 12.5% Score 136.6; DB 2; Length 2403;

Best Local Similarity 53.5%; Pred. No. 5.5e-27;

Matches 314; Conservative 0; Mismatches 264; Indels 9; Gaps 1;

272 GAACAGATCAACAACAATTCAGATTTATTTCCCAAGTATGAAGAAATGAGGCTT 331  
1756 GAGACCATGTGACCTCCATCTCGATGATGTTCCCAAGTATCTGGGACACACAGGCT 1815  
332 CCCATTAATTTGGGCTGCTGCTGTTGTTTCTTCTTCTGCTGCTGCTGCTGCTGCTG 391  
1816 GTGTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1875  
392 GCTGGAATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451  
1876 GGTGGAATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1935  
452 GCAGCTATACGAGCTAGTGAATCATCTGATTTATGAGGGAACAGATTCATTGAG 511  
1936 ATTGCATATTGAGCTTGTGGACATCTCTATGCTAGCGCTTGACAGAGGTTCTGTGAA 1995  
512 GATACAGAAATGATGTTGGACAAAGAGGTGATTTGCTATGAGGAGGCTGCTG 571  
1996 GACATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2046  
572 TGGTTGTAATTAACGCTATCTCTTGTGATTCGAATTTATCTGCTGCTGCTGCTGCTGCT 631  
2047 TGGGCGTTTGACACACGACATTTTACGTTTATCTTGTGCTGCTGCTGCTGCTGCTGCT 2106  
632 CATACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 691  
2107 GAGCCCATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2166  
692 ATGATGTTTTCGCTATTTTGGATTCGAATTTATCTGCTGCTGCTGCTGCTGCTGCTGCT 751  
2167 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2226  
752 AAAGAAACATCTTCAACGCTTATTAAGTTGCTGACACAGCTTCTAAGTGGGCTGCA 811  
2227 CCGGAGATTTATTAAGAGGCTGAAGTTGATGCTGCGCACAGCGGAGCTGGGCGCCA 2286  
812 TACCTGGAACAACATGCTGGGGAAGATTAAGAAAGATGATGATGCTGCTGCTGCTGCTG 858  
2287 TTCTTAGCTCAGACCGCGGGAACGCTACAGAAATATGATGACCC 2333

RESULT 13  
US-08-655-836-4

Sequence 4, Application US/08655836  
Patent No. 5824486

GENERAL INFORMATION:

APPLICANT: Borden, Laurence A.

APPLICANT: De Vivo, Michael

APPLICANT: Yokoyama, Midori

APPLICANT: Albert, Vivian R.

TITLE OF INVENTION: GLYCINE TRANSPORTER-TRANSFECTED CELLS

TITLE OF INVENTION: AND USES THEREOF

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dechert Price & Rhoads

STREET: P.O. Box 5218

CITY: Princeton

STATE: NJ

COUNTRY: USA

ZIP: 08543

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patenting Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/655,836

FILING DATE: 31-MAY-1996

CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:

NAME: Bloom, Allen

REGISTRATION NUMBER: 29135

REFERENCE/DOCKET NUMBER: 317743-105

TELECOMMUNICATION INFORMATION:

TELEPHONE: 609-520-3214

TELEPHONE: 609-520-3259

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 2817 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-655-836-4

Query Match 12.5% Score 136.6; DB 1; Length 2817;

Best Local Similarity 53.5%; Pred. No. 5.8e-27;

Matches 314; Conservative 0; Mismatches 264; Indels 9; Gaps 1;

272 GAACAGATCAACAACAATTCAGATTTATTTCCCAAGTATGAAGAAATGAGGCTT 331  
1963 GAGACCATGTGACCTCCATCTCGATGATGATGATGATGATGATGATGATGATGATGATG 2022  
332 CCCATTAATTTGGGCTGCTGCTGTTGTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 391  
2023 GTGTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2082  
392 GCTGGAATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451  
2083 GGTGGAATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2142  
452 GCAGCTATACGAGCTAGTGAATCATCTGATTTATGAGGGAACAGATTCATTGAG 511  
2143 ATTGCATATTGAGCTTGTGGACATCTCTATGCTAGCGCTTGCAAGAGGTTCTGTGAA 2202  
512 GATACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 571  
2203 GACATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2253  
572 TGGTTGTAATTAACGCTATCTCTTGTGATTCGAATTTATCTGCTGCTGCTGCTGCTGCTG 631  
2254 TGGGCGTTTGACACACGACATTTTACGTTTATCTTGTGCTGCTGCTGCTGCTGCTGCTG 2313  
2314 GAGCCCATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2373  
692 ATGATGTTTTCGCTATTTTGGATTCGAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 751  
2374 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2433  
752 AAAGAAACATCTTCAACGCTTATTAAGTTGCTGACACAGCTTCTAAGTGGGCTGCA 811  
2434 CCGGAGATTTATTAAGAGGCTGAAGTTGATGCTGCGCACAGCGGAGCTGGGCGCCA 2493  
812 TACCTGGAACAACATGCTGGGGAAGATTAAGAAAGATGATGATGCTGCTGCTGCTGCTG 858  
2494 TTCTTAGCTCAGACCGCGGGAACGCTACAGAAATATGATGACCC 2540



```

Db      2254  TGGGGCTTTTGTCACACCGACCAATTTAAACGTTTATCTTTCAGCTTCATCATGCG 2313
QY      632  CATAGACCTAAATTATGGCGCAATTCCATACCCTGATGGGAGTGTCTTTAGGGCGGT 691
Db      2314  GAGCCCATGACCTATGGGCTCTTACCGCTACCCCTAACTGTCATAGTGTCTTGATGGCTG 2373
QY      692  ATGATTGTTTCTTGCATTATTTGGATTTCCATTATATGCTATATCAATAAATAATTCAGCT 751
Db      2374  ATGCTGCGCTGCTCGTGATCTGGATCCGATTAATGATTCGATATAAAATGATATCTGGCT 2433
QY      752  AAGGAAACATCTTTCAAAGCCCTTATTAAGTTGCTGCACAGCCAGCTTCAACTGGGGTCCA 811
Db      2434  CCTGGAGATTTATTTAGAGAGGCTGAAGTTGATATGCTGCCACAGCCGAGACTGGGGCCCA 2493
QY      812  TACCTGSAACAACATCTGGGGGAAAGATATATAAGCATGTGATGTC 858
Db      2494  TTCTTAGCTACAGACCGCGGGGGAACGCTAACAAAGATATGATCGACC 2540

RESULT 15
US-09-191-468-89
; Sequence 89, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
; APPLICANT: Callagher, Michael J.
; APPLICANT: Burgess, Lloyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 123110S01
; CURRENT APPLICATION NUMBER: US/09/191.468A
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(450)
; OTHER INFORMATION: Seq 42(E)S2 - HSpC-2; nt 1942-2394; nt 1949 is A
; OTHER INFORMATION: not T (consensus); nt 1959 is C not T (consensus);
; US-09-191-468-89
; OTHER INFORMATION: nt 2130 is C not T (consensus)

```

Query Match	Best Local Similarity	Score	ID	Length
Matches	Conservative	Pred. No. 7.1e-18	Mismatches	Indels
9.38;	55.78;	101.4;	DB 4;	453;
220;	0;	166;	9;	Gaps
QY	464	GAGCTGTTGGATCATCTGATTTATTTGAGGGAAACAGATTTCATGTGAGATACAGAAATG	52	
DB	1	GACCTGAGAGGGATCTCCCTATGTGTATGTGCTTGCAAGATTTGTGAATATTAAAGATG	60	
QY	524	ATGATGGAGCAAGAGGTGATATTTTGGCTATGATGAGAGACTTGCTGGTTTGTAAT	58	
DB	61	ATGATTGGATTCACAGCTAACATCTTC-----TGAAGTGTGCTGGGCAATTGTA	11	
QY	584	ACGCTATCTCTTTGATTCGATATTTATCTGTATGTGGTCAATTTTCATAGACCTAAT	64	
DB	112	ACCCACACATTTTAACTTATCTCTTGCTTCACCTTTTACCACTGGGAACCCATGACC	17	
QY	644	TATGGCGCAATTCATACCCCTGACGTGGGAGTTGCTTATGAGCTGGTGTATGATTGTTTC	70	
DB	172	TATGGCTCTTACCGCTACCCCTAACGTCATAGTGTCGGATGGCTAATCTGCGCTGT	23	
QY	704	TGCATTATTTGGATTCCCAATTATGCTATTCATATAAAATATTCAGGCTAAAGAAATC	76	
DB	232	TCCGTCATCTGATGCCCAATATATGTTGTGTATATAAAATGATCTGCGCCCTGGAAAGATT	291	
QY	764	TTTCAAGCCCTTATATAGTTGCTGACAGACACACTTCACTGAGGGGTGCATACCTGGACAA	823	
DB	292	ATTGAGAGGCTGAAGTTGGTGTGCTGCGCACAGCCGACACTGAGGCCCATCTTATAGTCGAA	351	

Wed Aug 27 18:31:19 2003

us-09-940-227-15.rni

Page 10

QY 824 CATCGTGGGGAAGATATTAAGACATGTGTTCC 858  
||| ||||| ||| ||||| |||  
Db 352 CACCGGGGGAGCGTTACAGAACATGATCGACCC 386

Search completed: August 27, 2003, 12:08:56  
Job time : 94 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 11:20:56 : Search time 330 Seconds  
(without alignments)  
7605.789 Million cell updates/sec

Title: US-09-940-227-15

Perfect score: 1094  
Sequence: 1 gttcacaggagactgtacc.....tggtttttcaaaaaaaaaa 1094

Scoring table:  
IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCMS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09C\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	28.1	1232	13 US-10-027-632-101128	Sequence 101128,
2	307	28.1	1232	13 US-10-027-632-101129	Sequence 101129,
3	147.8	13.5	2863	9 US-09-795-232-1	Sequence 1, Appl1
4	65.4	6.0	553	10 US-09-878-178-1740	Sequence 1740, Ap
5	65.4	6.0	553	13 US-10-046-935-1740	Sequence 1740, Ap
6	65.4	6.0	553	14 US-10-146-503-1740	Sequence 1740, Ap
7	63.2	5.8	1800	12 US-10-345-680-6	Sequence 1740, Ap
8	63.2	5.8	2288	12 US-10-345-680-4	Sequence 4, Appl1
9	56.2	5.1	1911	12 US-10-289-161A-1	Sequence 4, Appl1
10	56.2	5.1	1911	12 US-10-289-161A-2	Sequence 2, Appl1
11	53.4	4.9	1938	12 US-09-845-908-3	Sequence 3, Appl1
12	50.4	4.6	2190	9 US-09-795-693-6	Sequence 6, Appl1
13	50.4	4.6	2190	14 US-10-156-239-6	Sequence 6, Appl1
14	50.4	4.6	2190	14 US-10-199-485-6	Sequence 6, Appl1
15	50.4	4.6	2204	9 US-09-741-149-1	Sequence 1, Appl1
16	50.4	4.6	2304	12 US-10-385-614-1	Sequence 1, Appl1

17	50.4	4.6	3103	9 US-09-795-693-4	Sequence 4, Appl1
18	50.4	4.6	3103	14 US-10-156-239-4	Sequence 4, Appl1
19	50.4	4.6	3103	14 US-10-199-485-4	Sequence 4, Appl1
20	50.2	4.6	424	10 US-09-960-352-7432	Sequence 7432, Ap
21	48.8	4.5	1991	12 US-09-845-908-9	Sequence 9, Appl1
22	47.8	4.4	2558	12 US-10-205-219-31	Sequence 31, Appl1
23	47.6	4.4	1833	10 US-09-861-846-1	Sequence 1, Appl1
24	45.4	4.1	3945	12 US-10-101-510-551	Sequence 551, App
25	45.4	4.1	3945	12 US-10-241-220-53	Sequence 53, App
26	45.4	4.1	3945	14 US-10-213-948-4	Sequence 5, Appl1
27	45	4.1	2093	12 US-09-845-908-5	Sequence 34, Appl1
28	45	4.1	3972	12 US-10-205-194-34	Sequence 950, App
29	44.6	4.1	18512	12 US-10-311-455-950	Sequence 1, Appl1
30	44.2	4.0	2860	12 US-10-298-974-1	Sequence 1, Appl1
31	44.2	4.0	2860	12 US-09-991-458-1	Sequence 1, Appl1
32	43.6	4.0	1051	12 US-09-845-908-7	Sequence 7, Appl1
33	43.6	4.0	2028	10 US-09-917-800A-1429	Sequence 1429, Ap
34	43.6	4.0	2028	12 US-09-845-908-1	Sequence 9, Appl1
35	43.6	4.0	4956	10 US-09-815-923-9	Sequence 444, App
36	43.2	3.9	6534	12 US-10-311-455-444	Sequence 487, App
37	42.8	3.9	6464	12 US-10-311-455-487	Sequence 1285, Ap
38	42.6	3.9	5858	12 US-10-311-455-1285	Sequence 1, Appl1
39	42.4	3.9	2485	9 US-09-923-444A-1	Sequence 1121, Ap
40	42.2	3.9	5452	12 US-10-311-455-1121	Sequence 1247, Ap
41	42.2	3.9	13038	12 US-10-311-455-1247	Sequence 14, Appl1
42	42	3.8	6113	12 US-10-204-708-14	Sequence 404, App
43	42	3.8	6113	12 US-10-311-455-404	Sequence 646, App
44	42	3.8	16287	12 US-10-311-455-646	Sequence 164, App
45	41.8	3.8	17137	12 US-10-311-455-164	

## ALIGNMENTS

RESULT 1  
US-10-027-632-101128  
Sequence 101128, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 101128  
LENGTH: 1232  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-101128

Query Match 28.1%; Score 307; DB 13; Length 1232;  
Best Local Similarity 96.9%; Pred. No. 2.1e-67;  
Matches 313; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 764 TTTCACGGCTATAGTTGTCGACACGCTTACTGCGCTCATCTGACAA 823  
DB 432 TTTCAGCGCTTATAGTTGTCGACACGCTTACTGCGCTCATCTGACAA 491

QY	824	CATCGTCGGGAAAAGATATTAAGACATGGTGTGCTCTAAAAAGAGCGCCATGAATA	883
QY	492	CATCGTGGGAAAATATTAAGACATGGTACATCTTAAAAAGAGCGCTGACCATGAATA	551
QY	884	CCTACTGTTAGTGGCAGCAGAAAACCGGAATGAGATCTCATTTGAAAAAATATATGATTG	943
Db	552	CCCTACTGTAGTGGCAGCAGAAAACCGGAATGAGATCTCATTTGAAAAAATATATGATTG	611
QY	944	TATATGTGATTTTTTTTAGAATGAGGGGACCCGTATTTATTTGCTGTAAAGAAATAG	10030
Db	612	TATATGTGATTTTTTTTAGAATGAGGGGAACTTATTTATTTGCTGTAAAGAAATAG	671
QY	1004	GAATATGTACTACTATGTCATGATGAGGTGATTTTTTTTCCATTTAAGCAGGATGCA	1063
Db	672	GAATATGTACTACTATGTCATGATGATGATGATTTTTTTTCCATTTAAGCAGGATGCA	731
QY	1064	ATATATAAATGTGTTTTTTAA	1086
Db	732	ATATATAAATGTGATCTCTTAA	754

	RESULT 2	
DB	US-10-027-632-101129	
	; Sequence 101129, Application US/10027632	
	; GENERAL INFORMATION:	
	; APPLICANT: Wang, David G.	
	; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide	
	; TITLE OF INVENTION: Polymorphisms in the Human Genome	
	; FILE REFERENCE: 108827.129	
	; CURRENT APPLICATION NUMBER: US/10/027.632	
	; CURRENT FILING DATE: 2002-04-30	
	; PRIOR APPLICATION NUMBER: US 60/218,006	
	; PRIOR FILING DATE: 2000-07-12	
	; PRIOR APPLICATION NUMBER: US 60/198,676	
	; PRIOR FILING DATE: 2000-04-20	
	; PRIOR APPLICATION NUMBER: US 60/193,483	
	; PRIOR FILING DATE: 2000-03-29	
	; PRIOR APPLICATION NUMBER: US 60/185,218	
	; PRIOR FILING DATE: 2000-02-24	
	; PRIOR APPLICATION NUMBER: US 60/167,363	
	; PRIOR FILING DATE: 1999-11-23	
	; PRIOR APPLICATION NUMBER: US 60/156,358	
	; PRIOR FILING DATE: 1999-09-28	
	; PRIOR APPLICATION NUMBER: US 60/146,002	
	; PRIOR FILING DATE: 1999-08-09	
	; NUMBER OF SEQ ID NOS: 325720	
	; SOFTWARE: FastSeq for Windows Version 4.0	
	; SEQ ID NO 101129	
	; LENGTH: 1232	
	; TYPE: DNA	
	; ORGANISM: Human	
	US-10-027-632-101129	
	Query Match	28.1%; Score 307; DB 13; Length 1232;
	Best Local Similarity	96.9%; Pred. No. 2.1e-67;
	Matches 313; Conservative	0; Mismatches 10; Indels 0; Gaps 0;
OY		
	764 TTTCAGCCCTTATAGTTGCGTGACGACCCTTACTAGGGGGCCATACCTGGAAACA	823
DB	432 TTTTCGCCCTTATATAGTTGCTGCACGACCTTTTACTGGGGTCCATCCTGGAAACA	491
OY		
	824 CATCGTGGGGAAGAATATAAAGACATGTAGTCTCTAAAAAAGAGCGTGGCATGAATA	883
DB	492 CATCGTGGGGAAGAATATAAAGACATGTAGTCTCTAAAAAAGAGCGTGGCATGAATA	551
OY		
	884 CCTACTGTTTAGTGCAGCAAGAAAACCGGAATGAGATCTCATTTGAATAAATATATGATTG	943
DB	552 CCTACTGTTTAGTGCAGCAAGAAAACCGGAATGAGATCTCATTTGAATAAATATATGATTG	611
OY		
	944 TATATATGATTTTTTTTATAGATATAGGGGACCCCTATATTTTGTCTTAACGTAATG	1003
DB	612 TATATATGATTTTTTTTATAGATATAGGGGACCCCTATATTTTGTCTTAACGTAATG	671

Accession	Sequence	Position
QY	1004 GAAATGTACAACTACTATGTCATCATATGAGGATTTTTCCTCCATTTAAAGCAGAATGCA	1063
Db	672 GAAATGTACAACTACTATGTCATCATATGTCATTTTTCACATTTAAAGCAGAATGCA	731
QY	1064 ATATATAAAGTGGTTTTTTAA	1086
Db	732 ATATATAAAGTGAATCTCTTAA	754

Query Match	Best Local Similarity	Score	DB	Length	2863;
Matches 321; Conservative	54.7%;	Pred. No. 6.3e-27;	Indels 9;	Gaps 1;	
US-09-795-232-1					
Sequence 1, Application US/09795232					
Patent No. US20010012627A1					
GENERAL INFORMATION:					
APPLICANT: Anthony M. Brown					
APPLICANT: Conrad Gerald Chapman					
APPLICANT: Israel Simon Gloger					
APPLICANT: Joanne Rachel Evans					
APPLICANT: William Cairns					
APPLICANT: Hugh Jonathan Herdon					
TITLE OF INVENTION: NOVEL COMPOUNDS					
FILE REFERENCE: GP-30176-D1					
CURRENT APPLICATION NUMBER: US/09/795,232					
CURRENT FILING DATE: 2001-02-28					
PRIOR APPLICATION NUMBER: 09/182,728					
PRIOR FILING DATE: 1998-10-29					
PRIOR APPLICATION NUMBER: 9818890.7					
PRIOR FILING DATE: 1998-08-28					
NUMBER OF SEQ ID NOS: 6					
SOFTWARE: FastSeq for Windows Version 3.0					
SEQ ID NO 1					
LENGTH: 2863					
TYPE: DNA					
ORGANISM: HOMO SAPIENS					
US-09-795-232-1					
Query Match	13.5%;	Score 147.8;	DB 9;	Length 2863;	
Best Local Similarity	54.7%;	Pred. No. 6.3e-27;	Indels 9;	Gaps 1;	
Matches 321; Conservative	0;	Mismatches 257;	Indels 9;	Gaps 1;	
US-09-795-232-1					
Sequence 1, Application US/09795232					
Patent No. US20010012627A1					
GENERAL INFORMATION:					
APPLICANT: Anthony M. Brown					
APPLICANT: Conrad Gerald Chapman					
APPLICANT: Israel Simon Gloger					
APPLICANT: Joanne Rachel Evans					
APPLICANT: William Cairns					
APPLICANT: Hugh Jonathan Herdon					
TITLE OF INVENTION: NOVEL COMPOUNDS					
FILE REFERENCE: GP-30176-D1					
CURRENT APPLICATION NUMBER: US/09/795,232					
CURRENT FILING DATE: 2001-02-28					
PRIOR APPLICATION NUMBER: 09/182,728					
PRIOR FILING DATE: 1998-10-29					
PRIOR APPLICATION NUMBER: 9818890.7					
PRIOR FILING DATE: 1998-08-28					
NUMBER OF SEQ ID NOS: 6					
SOFTWARE: FastSeq for Windows Version 3.0					
SEQ ID NO 1					
LENGTH: 2863					
TYPE: DNA					
ORGANISM: HOMO SAPIENS					
US-09-795-232-1					

QY 752 AAGGAACATCTTTCACAGCCTTATAAGTTGCTGACAGACGCTTAACTGGGTC 811  
Db 2476 CCTGAGAGATTATTGAGAGAGCTGAGTGGCTGCCAGCGGACTGGGCCCA 2535  
QY 812 TACTGGAAACATCTCGTGGGGAAGAATATAAGCATGCTAGTCC 858  
Db 2536 TTCTTACCTCAACACCGGGGACGCTTACAAAGAATGATGACCC 2582

RESULT 4  
US-09-878-178-1740  
: Sequence 1740, Application US/09878178  
: Patent No. US2002017752A1  
: GENERAL INFORMATION:  
: APPLICANT: Jiang, Yugu  
: APPLICANT: Harlocker, Susan L.  
: APPLICANT: Secretist, Heather  
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
: FILE REFERENCE: 210121.527  
: CURRENT APPLICATION NUMBER: US/09/878.178  
: CURRENT FILING DATE: 2001-06-08  
: NUMBER OF SEQ ID NOS: 2237  
: SOFTWARE: FASTSEQ for Windows Version 4.0  
: SEQ ID NO 1740  
: LENGTH: 553  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: misc\_feature  
: LOCATION: (1)...(553)  
: OTHER INFORMATION: n = A,T,C or G  
US-09-878-178-1740

Query Match  
Best Local Similarity 92.0%; Score 65.4; DB 10; Length 553;  
Matches 69; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1012 ACATACATGTTTCATGATAGGATGATTTTCCCATTTAAGCAGATGCAATATAAA 1071  
Db 1 ACATACATGTTTCATGATAGGATGATTTTTCACATTTAAGCAGATGCAATATAAA 60

QY 1072 ATGTGCTTTTAA 1086  
Db 61 ATGTGAATCTCTTAA 75

RESULT 5  
US-10-046-935-1740  
: Sequence 1740, Application US/10046935  
: Publication No. US2002015601A1  
: GENERAL INFORMATION:  
: APPLICANT: Jiang, Yugu  
: APPLICANT: Harlocker, Susan L.  
: APPLICANT: Secretist, Heather  
: APPLICANT: Wang, Aijun  
: APPLICANT: Stolk, John A.  
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
: FILE REFERENCE: 210121.527C1  
: CURRENT APPLICATION NUMBER: US/10/046.935  
: CURRENT FILING DATE: 2002-01-15  
: NUMBER OF SEQ ID NOS: 2239  
: SOFTWARE: FASTSEQ for Windows Version 4.0  
: SEQ ID NO 1740  
: LENGTH: 553  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: misc\_feature  
: LOCATION: 546  
: OTHER INFORMATION: n = A,T,C or G  
US-10-046-935-1740

QY 1012 ACATACATGTTTCATGATAGGATGATTTTCCCATTTAAGCAGATGCAATATAAA 1071  
Db 1 ACATACATGTTTCATGATAGGATGATTTTTCACATTTAAGCAGATGCAATATAAA 60

QY 1072 ATGTGCTTTTAA 1086  
Db 61 ATGTGAATCTCTTAA 75

RESULT 6  
US-10-146-502-1740  
: Sequence 1740, Application US/10146502  
: Publication No. US20030069180A1  
: GENERAL INFORMATION:  
: APPLICANT: Jiang, Yugu  
: APPLICANT: Harlocker, Susan L.  
: APPLICANT: Secretist, Heather  
: APPLICANT: Wang, Aijun  
: APPLICANT: Stolk, John A.  
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
: FILE REFERENCE: 210121.527C2  
: CURRENT APPLICATION NUMBER: US/10/146.502  
: CURRENT FILING DATE: 2002-05-14  
: NUMBER OF SEQ ID NOS: 2241  
: SOFTWARE: FASTSEQ for Windows Version 4.0  
: SEQ ID NO 1740  
: LENGTH: 553  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: misc\_feature  
: LOCATION: 546  
: OTHER INFORMATION: n = A,T,C or G  
US-10-146-502-1740

Query Match  
Best Local Similarity 92.0%; Score 65.4; DB 14; Length 553;  
Matches 69; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1012 ACATACATGTTTCATGATAGGATGATTTTCCCATTTAAGCAGATGCAATATAAA 1071  
Db 1 ACATACATGTTTCATGATAGGATGATTTTTCACATTTAAGCAGATGCAATATAAA 60

QY 1072 ATGTGCTTTTAA 1086  
Db 61 ATGTGAATCTCTTAA 75

RESULT 7  
US-10-345-680-6  
: Sequence 6, Application US/10345680  
: Publication No. US20030148394A1  
: GENERAL INFORMATION:  
: APPLICANT: Millennium Pharmaceuticals, Inc.  
: APPLICANT: Sinos-Santiago, Inmaculada  
: APPLICANT: Venkateswarlu, Karicheti  
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
: TITLE OF INVENTION: UROLOGICAL DISORDERS USING 1435, 559, 34021, 44099, 25278,  
: TITLE OF INVENTION: 641, 260, 55089, 21407, 42032, 46656, 62553, 302, 323,  
: TITLE OF INVENTION: 12303, 985, 13237, 13601, 18926, 318, 2058 OR 6351 MOLECULES.  
: FILE REFERENCE: MPI02-012P1RM.OMN  
: CURRENT APPLICATION NUMBER: US/10/345.680  
: CURRENT FILING DATE: 2003-01-16  
: PRIOR APPLICATION NUMBER: US 60/349,511  
: PRIOR FILING DATE: 2002-01-18  
: PRIOR APPLICATION NUMBER: US 60/360,500  
: PRIOR FILING DATE: 2002-02-28

```
; PRIOR APPLICATION NUMBER: US 60/365,041
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/374,063
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/403,468
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 60/414,262
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US 60/419,986
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/423,809
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/429,797
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1800)
US-10-345-680-6
```

```
Query Match          5.8%; Score 63.2; DB 12; Length 1800;
Best Local Similarity 49.4%; Pred. No. 1.2e-05;
Matches 229; Conservative 0; Mismatches 223; Indels 12; Gaps 2;
```

```
QY 260 TCCCCCTCTGAGAAAGATACACACAACTTCAAGATTATTTCCCAAGTGAAG 319
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1192 TTTCGACCTGAGGGCTTCATCAGACGCCCTGGTGAATGAGTACCCAGCTCCCGC 1251
QY 320 AAAATGAGGTTCCCATACCTTTGGGCTGCTGCTGTTTGTCTTCTGCTGCTC 379
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1252 AACCGAGAGAGCTCTTCATTCGTCGTGCATATTCCTACCTGATCGGCTCTCT 1311
QY 380 TGTGTGCTGAGGCTGGAATTTACTGGTTCATCTGATGATACCATCTGTGTC---TGA 436
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1312 AACATCCTCAGGGGGATTTATGCTTCAAACTCTTACACTACTCTGCGACAGGCG 1371
QY 437 TGGGGCATTTTAATTCAGAGCTATACAGCTAGTGGAAATCATCTGGATTATGAGGG 496
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1372 ATGAGCCTGCTGTCTCCGTCGTCTTGAATGTCCTATTCCTGTTTACGGGTGTC 1431
QY 497 AACAGATTCATTAGAGTATACAGAAATGATGATGGACCAAGAGTGATATTCGCTA 556
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1432 AACCGATTCTATGACATATCCAGAGATGTTGGATCCAGGCCCTGCATC----- 1482
QY 557 TGGTGAGAGCTTGTGTTTGAATTAACGCTATCCTTTTGAATTGCAATTTATCTGG 616
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1483 TGGTGGAACTCTGCTGCTGCTTTCTTTCACACCAATCTTGGGGGCGGTTCATTTTC 1542
QY 617 TCATTGTCGAATTTTCATAGACCTAATTAATGCGCAATTCACCTGACTGGGAGTT 676
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1543 AGTGTGTGTGAGATGAGCCACACCATGGAATATGTTTCCCAAGTGGGGCAG 1602
QY 677 GCTTTAGCTGTGTATGATTTGTTTCTGCATTTATTTGATTC 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1603 GGTTGTGGCTGTGCTGATGCTCTGTCTTCATGCTCTCATCC 1646
```

## RESULT 8

```
US-10-345-680-4
; Sequence 4, Application US/10345680
; Publication No. US20030148394A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Sinos-Santiago, Immaculada
; APPLICANT: Venkateswarlu, Karicheti
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 1435, 559, 34021, 44099, 25278,
; TITLE OF INVENTION: 641, 260, 55089, 21407, 42032, 46656, 62553, 302, 323,
```

```
; TITLE OF INVENTION: 12303, 985, 13237, 13601, 18926, 318, 2058 OR 6351 MOLECULES.
; FILE REFERENCE: MP102-012PIRNM.OMI
; CURRENT APPLICATION NUMBER: US/10/345,680
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US 60/349,511
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/360,500
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/365,041
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/374,063
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/403,468
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 60/414,262
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US 60/419,986
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/423,809
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/429,797
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (235)...(2034)
US-10-345-680-4
```

```
Query Match          5.8%; Score 63.2; DB 12; Length 2298;
Best Local Similarity 49.4%; Pred. No. 1.4e-05;
Matches 229; Conservative 0; Mismatches 223; Indels 12; Gaps 2;
```

```
QY 260 TCCCCCTCTGAGAAAGATACACACAACTTCAAGATTATTTCCCAAGTGAAG 319
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1426 TTTCGACCTGAGGGCTTCATCAGACGCCCTGGTGAATGAGTACCCAGGCTCCCGC 1485
QY 320 AAAATGAGGTTCCCATACCTTTGGGCTGCTGCTGTTTGTCTTCTGCTGCTC 379
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1486 AACCGAGAGAGCTCTTCATTCGTCGTGCATATTCCTACCTGATCGGCTCTCT 1545
QY 380 TGTGTGCTGAGGCTGGAATTTACTGGTTCATCTGATGATACCATCTGTGTC---TGA 436
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1546 AACATCCTCAGGGGGATTTATGCTTCAAACTCTTGACTACTACTGCGACATGCG 1605
QY 437 TGGGGCATTTTAATTCAGAGCTATACAGCTAGTGGAAATCATCTGGATTATGAGGG 496
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1606 ATGAGCCTGCTGTCTCCGTCGTCTTGAATGTCCTATTCCTGTTTACGGGTGTC 1665
QY 497 AACAGATTCATTAGAGTATACAGAAATGATGATGGACCAAGAGTGATATTCGCTA 556
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1666 AACCGATTCTATGACATATCCAGAGATGTTGGATCCAGGCCCTGCATC----- 1716
QY 557 TGGTGAGAGCTTGTGTTTGAATTAACGCTATCCTTTTGAATTGCAATTTATCTGG 616
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1717 TGGTGGAACTCTGCTGCTGCTTTCTTTCACACCAATCTTGGGGGCGGTTCATTTTC 1776
QY 617 TCATTGTCGAATTTTCATAGACCTAATTAATGCGCAATTCACCTGACTGGGAGTT 676
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1777 AGTGTGTGTGAGATGAGCCACACCATGGAATATGTTTCCCAAGTGGGGCAG 1836
QY 677 GCTTTAGCTGTGTATGATTTGTTTCTGCATTTATTTGATTC 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1837 GGTTGTGGCTGTGCTGATGCTCTGTCTTCATGCTCTCATCC 1880
```

## RESULT 9

```
US-10-289-161A-1
; Sequence 1, Application US/10289161A
```

```
;; Publication No. US20030152970A1
;; GENERAL INFORMATION:
;; APPLICANT: Millennium Pharmaceuticals, Inc.
;; APPLICANT: Sinos-Santago, Immaculada
;; TITLE OF INVENTION: Methods and Compositions to Treat Pain
;; TITLE OF INVENTION: and Painful Disorders Using 577, 20739 or 57145
;; FILE REFERENCE: MPI2001-287PIR(M)
;; CURRENT APPLICATION NUMBER: US/10/289,161A
;; PRIOR FILING DATE: 2003-03-25
;; PRIOR APPLICATION NUMBER: US 60/333,073
;; PRIOR FILING DATE: 2001-11-06
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1
;; LENGTH: 1911
;; TYPE: DNA
;; ORGANISM: Homo Sapiens
US-10-289-161A-1
```

```
Query Match 5.18; Score 56.2; DB 12; Length 1911;
Best Local Similarity 44.9%; Pred. No. 0.00076;
```

```
Matches 266; Conservative 0; Mismatches 318; Indels 9; Gaps 1;
```

```

OY 264 CCTGTGAAGAAAGATCAACAACATTTCAAGATTTATTTCCCAAGTATGAAGAAA 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1205 CCTTCTGAGAGACCATTTGTGACAGCTGTGACAGATGAGTTCCCATTAACCTGGGCCCA 1264
OY 324 TGAGGGTTCCTTAACCTTTGGGCTGCTGCTTTGTTTCTTCCTTGGTCTGCTGTG 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1265 AGAAGGGGCTGTTCTCAGGGCTCATCTGCGGGCATCTACATGATGGGCTGATCTCA 1324
OY 384 TGACTAGGCTGGAATTTACTGGGCTTATCTGATTTGACACATCTGTGCTGATGGGCA 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1325 CCACTATGAGGGGATCTACTGCTGCTGCTTCTGATGACTACAGCGCCAGCTTGGGC 1384
OY 444 TTTTAATTGAGCTATACGTGAGCTAGTTGGAATCAATCAATTTATGAGGAAACAT 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1385 TGATGCTGTGCTTATACCAACATGCTTCCCTGACACGGGTGTATGCAATTCAGAGGT 1444
OY 504 TCATTGAGATACAGAAATGATTTGAGCAAAAGAGTGATTTCTGCTATGCTGGA 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1445 TCTGCCAGACATCCACATGATGCTGGGCTTCAAGCGGGGCTCTACTTC-----A 1495
OY 564 GAGCTGCTGCTTTGTAATTAACGCTTATCTTTGATTCGAATTTATCTGCTCATTTGG 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1496 GGGCTGCTGCTGCTTCTCTCCCAAGCCAGCTCTTGGCCCTCTGCTGTATAGCATCG 1555
OY 624 TGCATTTTCAATAGACCTAATTTATGAGGCAATTTCCATPACCTGACTGGGGGCTTTAG 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1556 TCAAGTACCAAGCCCTCGAGATATGCGATTAACGCTTCCCGCCCTGGGCTGAGCTGTGG 1615
OY 684 GCTGTGTATGATTTTCTGCAATTTTGAATTCATTAATGCTATCAATAAATAA 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1616 GCATCTGTATGAGGCTGCTGCTCTGCTCATGATCCAGCTGAGCTGTGCTGTGC 1675
OY 744 TTCAGGCTTAAGAAACATCTTTCAAGCCTTTAATAGTTGCTGCAACACGCTTCTACT 803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1676 TTCGAAAGAGGGCTCACTCTGGAGCGGCTCCACAGGCGACCGCGGCATGAGACT 1735
OY 804 GGGGTCATACCTGGAACAACATCTGGGGAAGATATAAGCATGAGTGT 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1736 GGGGACCATCTGCGAGAGAAACCGAGCGGCTATGTGTGGCCAGCTGTGCT 1788
```

```
RESULT 10
US-10-289-161A-2
; Sequence 2, Application US/10289161A
; Publication No. US20030152970A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Sinos-Santago, Immaculada
; TITLE OF INVENTION: Methods and Compositions to Treat Pain
; TITLE OF INVENTION: and Painful Disorders Using 577, 20739 or 57145
```

```
;; FILE REFERENCE: MPI2001-287PIR(M)
;; CURRENT APPLICATION NUMBER: US/10/289,161A
;; CURRENT FILING DATE: 2003-03-25
;; PRIOR APPLICATION NUMBER: US 60/333,073
;; PRIOR FILING DATE: 2001-11-06
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 1911
;; TYPE: DNA
;; ORGANISM: Homo Sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)...(1911)
US-10-289-161A-2
```

```
Query Match 5.18; Score 56.2; DB 12; Length 1911;
Best Local Similarity 44.9%; Pred. No. 0.00076;
```

```
Matches 266; Conservative 0; Mismatches 318; Indels 9; Gaps 1;
```

```

OY 264 CCTGTGAAGAAAGATCAACAACATTTCAAGATTTATTTCCCAAGTATGAAGAAA 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1205 CCTTCTGAGAGACCATTTGTGACAGCTGTGACAGATGAGTTCCCATTAACCTGGGCCCA 1264
OY 324 TGAGGGTTCCTTAACCTTTGGGCTGCTGCTTTGTTTCTTCCTTGGTCTGCTGTG 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1265 AGAAGGGGCTGTTCTCAGGGCTCATCTGCGGGCATCTACATGATGGGCTGATCTCA 1324
OY 384 TGACTAGGCTGGAATTTACTGGGCTTATCTGATTTGACACATCTGTGCTGATGGGCA 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1325 CCACTATGAGGGGATCTACTGCTGCTGCTTCTGATGACTACAGCGCCAGCTTGGGC 1384
OY 444 TTTTAATTGAGCTATACGTGAGCTAGTTGGAATCAATCAATTTATGAGGAAACAT 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1385 TGATGCTGTGCTTATACCAACATGCTTCCCTGACAGGGGTATGCAATTCAGAGGT 1444
OY 504 TCATTGAGATACAGAAATGATTTGAGCAAAAGAGTGATTTCTGCTATGCTGGA 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1445 TCTGCCAGACATCCACATGATGCTGGGCTTCAAGCGGGGCTCTACTTC-----A 1495
OY 564 GAGCTGCTGCTTTGTAATTAACGCTTATCTTTGATTCGAATTTATCTGCTCATTTGG 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1496 GGGCTGCTGCTGCTTCTCTCCCAAGCCAGCTCTTGGCCCTCTGCTGTATAGCATCG 1555
OY 624 TGCATTTTCAATAGACCTAATTTATGAGGCAATTTCCATPACCTGACTGGGGGCTTTAG 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1556 TCAAGTACCAAGCCCTCGAGATATGCGATTAACGCTTCCCGCCCTGGGCTGAGCTGTGG 1615
OY 684 GCTGTGTATGATTTTCTGCAATTTTGAATTCATTAATGCTATCAATAAATAA 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1616 GCATCTGTATGAGGCTGCTGCTCTGCTCATGATCCAGCTGAGCTGTGCTGTGC 1675
OY 744 TTCAGGCTTAAGAAACATCTTTCAAGCCTTTAATAGTTGCTGCAACACGCTTCTACT 803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1676 TTCGAAAGAGGGCTCACTCTGGAGCGGCTCCACAGGCGACCGCGGCATGAGACT 1735
OY 804 GGGGTCATACCTGGAACAACATCTGGGGAAGATATAAGCATGAGTGT 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1736 GGGGACCATCTGCGAGAGAAACCGAGCGGCTATGTGTGGCCAGCTGTGCT 1788
```

```
RESULT 11
US-09-845-908-3
; Sequence 3, Application US/09845908
; Publication No. US20030143729A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Kell E. et al
; TITLE OF INVENTION: DNA Encoding Taurine and GABA Transporters and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 40558-D
; CURRENT APPLICATION NUMBER: US/09/845,908
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 28
```





Query Match	4.68;	Score 50.4;	DB 14;	Length 2190;
Best Local Similarity	53.68;	Pred. No. 0.024;		
Matches 105; Conservative	0;	Mismatches 91;	Indels 0;	

RESULT 15  
US-09-741-149-1

Query Match	4.68;	Score 50.4;	DB 9;	Length 2204;
Best Local Similarity	53.68;	Pred. No. 0.024;		
Matches 105; Conservative	0;	Mismatches 01;		

Oy 336 TAACCTGGGCGTGCCTGGTTGATTTCCTGTGCTCAGTGAGACTACAGGCCG 395  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 1490 TATCTATTAATGCTGTCTCTCGCATTTATTAATGAGCTGAATTTGGCAACGCTCG 1548  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Oy 396 GAATTACTGGGTTATCTGATTCACCCTCTGTGCTGGATGGGCATTTAATTGAC 455  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 1550 GAAATTACTTTGTTACAAATGTTTGCATGATTAATTCGTACACTGCCTTCGTAATTGG 1609

OY 456 CTATACGAGACCTAGTTGGAAATCATCTGGATTTATGAGGGACAGATTTCATTGAGATA 515

Db 1610 TCATTTTGGAATAATTCCTCTGTATGCTTTGTTTATGCGATGAAATAGTTTATGGAAGACC 1669

OY 516 CAGAAATGATGATTTGG 531

Db 1670 TAAAGATATGCTGGG 1685

```
Search completed: August 27, 2003, 17:32:36
Job time : 334 secs
```

**THIS PAGE BLANK (USPTO)**

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 09:55:34 ; Search time 2773 Seconds  
(without alignments)  
9588.567 Million cell updates/sec

Title: US-09-940-227-15

Perfect score: 1094  
Sequence: 1 gttcacaggaggactgtacc.....tggttttttaaaaaaaaaa 1094

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estb9:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrt:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	621.8	56.8	665	13	BU687102 UI-CF-DU1
C 2	611.8	55.9	695	9	AM190954
C 3	561	51.3	2199	11	AK018553
C 4	532.8	48.7	574	9	AI669617

Result No.	Score	Query Match	Length DB	ID	Description
C 5	528	48.3	617	9	AM812970
C 6	507	46.3	530	12	BM826722
C 7	476.6	43.6	691	12	BG964290
C 8	472.6	43.2	487	9	AA552658
C 9	468.8	42.9	501	10	BF096011
C 10	425.2	38.9	614	14	CB579897
C 11	415	37.9	567	4	BX523178
C 12	411.6	37.6	582	14	CB586665
C 13	381	34.8	484	10	BF774564
C 14	344.2	31.5	438	9	AI429024
C 15	325.8	29.8	360	10	BE003214
C 16	324	29.6	460	13	BQ378904
C 17	290	26.5	332	10	BE002150
C 18	276	25.2	463	10	AI605513
C 19	273.8	25.0	386	9	AA592728
C 20	267.6	24.5	682	9	AL871486
C 21	246.6	22.5	636	9	AL893341
C 22	222	20.3	893	13	BU899741
C 23	219.4	20.1	889	14	CA791412
C 24	218	19.9	374	10	BE819593
C 25	213	19.5	337	13	BQ082831
C 26	209	19.1	498	9	AL872634
C 27	207.4	19.0	482	13	BQ311137
C 28	205.6	18.8	282	10	BE183155
C 29	205.6	16.1	201	13	BQ378913
C 30	176.4	11.5	1068	13	BM559743
C 31	125.4	11.3	677	13	BM302794
C 32	124	11.0	740	14	CA312866
C 33	120	9.5	142	10	BE005236
C 34	104.4	8.1	574	9	AM645092
C 35	101.8	8.1	574	9	AM645092
C 36	88.4	8.0	553	10	BG467430
C 37	87.4	7.6	345	13	BX095910
C 38	83	7.5	151	10	BE814811
C 39	82	7.3	660	13	BM045601
C 40	80.2	7.2	806	13	BM017451
C 41	78.8	7.0	641	13	BU140319
C 42	77	7.0	787	12	BG781704
C 43	76.4	6.8	757	13	BU141486
C 44	74.2	6.6	234	10	BU141317
C 45	72.2				

## ALIGNMENTS

RESULT 1  
BU687102 665 bp mRNA linear EST 07-OCT-2002  
UI-CF-DU1-ado-b-18-0-UI-s1 UI-CF-DU1 Homo sapiens CDNA clone  
DEFINITION  
UI-CF-DU1-ado-b-18-0-UI 3', mRNA sequence.  
ACCESSION  
BU687102  
VERSION  
BU687102.1 GI:23542655  
KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

REFERENCE  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 665)

AUTHORS  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
TITLE  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL  
Genome Res. 6 (9), 791-806 (1996)  
MEDLINE  
97044477  
PUBMED  
8889548

COMMENT  
Contract: McCray, PB

McCray Lab  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: pauli-mccray@iowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa





```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:9030613J17"
/db_xref="MGI:1907717"
/db_xref="taxon:10090"
/clone="9030613J17"
/sex="male"
/tissue_type="colon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
108..2024
/Note="unnamed protein product; NA+ AND CL-COUPLED NEUTRAL
AND BASIC AMINO ACID TRANSPORTER AIBO,+ (SPR1)09160,
evidence: FASTV, 99.7%ID, 100%length, match=1914)
putative"
/codon_start=1
/protein_id="BAB31272.1"
/db_xref="GI:12858314"
/db_xref="MGI:1890216"
/translation="MDRLCPNPFKCRKQKRYTASSENHVEGENDENRGNMKSXD
YLLSWGVAVGLGNWREPLYLFTNGGAFILIPYAIMLAGLPLFLECSLQGFSL
GPRVWRILPLFOGCTIMVLSVFAIYVNIYLSYLFASQSVLPWANSMA
DENCSRPLVTCGNVSTGNGEPMNISMVNTNLTGNGSEVEREPGALPSROYMDKT
LORSSGMDTGVVWTLACLILAMILVGAALFKIKSKGKVVETLPPVLIIL
IRGATLEGASKGISYIGASNFTKIREAVKDAOTFLISLVANGGVALSYNK
FNNCYSDAIIVCLNCLTSVFAIFSLI6HMAHISKEVSPQVKSQFDLATAP
BALAOLPAPFWSILFFMLTLGLDSQFASLETITTFQDLFFPAMKRMVPTLGC
CLIFLILGLCVQAGIYVWHLIDHSCAGMGLIALEIAGIIVYGNRPITLDEM
MIGARKWTFMIRACWFIYTPILSALVMSVFERPDYADIEPYDGVALGCMCI
IFCLINPIMAIKTIQVQBNILQRIISCIPTASNMGPYLEKHGRERYDMAPEAKET
DHEIPISGSTPE"
2181..2186
polyA_signal
/Note="putative"
2199
polyA_site
/Note="putative"
BASE COUNT 590 a 404 c 507 g 698 t
ORIGIN
Query Match 51.3%; Score 561; DB 11; Length 2199;
Best Local Similarity 81.6%; Pred. No. 9.3e-110;
Matches 675; Conservative 0; Mismatches 145; Indels 7; Gaps 2;

```

```

Db 1792 GCTGCTGATGATCAATTTCTGCAATATTGGATACCAATATAGGCTATCATAAATAG 1851
Qy 744 TTCAGCTAAAGAAACATCTTTCACAGCCTATATAGTGGCGAGACCACTTCTAAT 803
Db 1852 TTCAGCGCAAGAGAAACATCTTTCACAGCCTATATAGTGGCGAGACCACTTCTAAT 1911
Qy 804 GGGGTCATACCTCGAACAACATCGTGGGAAAGATATTAAGACATGTAGTCTTA 863
Db 1912 GGGGTCATACCTCGAACAACATCGTGGGAGAGATACAGACATGTAGTCTTA 1971
Qy 864 AAGAGCTGGCCATGAATACCTACTGTATAGGCGAGAGAAACCGAATGATCTCA 923
Db 1972 AGGAGACTGACCAAGAAATACCTACTGTATAGGCGAGAGAAACCGAATGATCTCA 2031
Qy 924 ---TTGAAAAAATATATGATGTATATATGATATTTTATAGATAGGGGACCTTAT 980
Db 2032 CTTTGAATAATATCTAATTTGTATATATGATTTTGTAGAGTAAAGGATTTTAT 2091
Qy 981 TTTATTTGTGTATACGAATAGAAATGTACATCTATGTCATGATAGGT---GA 1036
Db 2092 TTATTTGTATGATTAATTAATAGGAATGTACATGTCATGTCATGATAGGT 2151
Qy 1037 TTTTTCCTCATTAACGAGAAATGCAATATAAATGTGTTT 1083
Db 2152 TACTTTTTCATTAATCAAGATGAATATTAATATGATCTAT 2198

RESULT 4
A1669617/c 574 bp mRNA linear EST 15-DEC-1999
LOCUS tw34e04.x1 NCI-CGAP Utl Homo sapiens cDNA clone IMAGE:2261598.3
DEFINITION similar to TR:015003 015003 GUT1-1 LIKE.; mRNA sequence.
ACCESSION A1669617
VERSION A1669617
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 574)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
JOURNAL COMMENT
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/UM/NIH at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 1364 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 385.
Location/Qualifiers
1..574
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2261598"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/ab_host="DH10B"
/clone_lib="NCI-CGAP-Utl"
/Note="Organ: uterus; Vector: pCMV-Sport6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"
BASE COUNT 182 a 118 c 80 g 193 t 1 others
ORIGIN

```

Query Match 48.7%; Score 532.8; DB 9; Length 574;  
 Best Local Similarity 99.3%; Pred. No. 9.3e-104;  
 Matches 545; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Oy 546 TATTCGCTATGCTGAGACCTGCTGTTTGTAAATTAACGCTATCCCTTTGATGCA 605  
 Db 573 TATTCGCTATGCTGAGAGCTGCTGCTG- TTGTAATTAACGCTATATCCTTTGATGCA 515  
 Oy 606 TATTAATGCTATGCTGCTGCAATTTATGACCTAATTTGGCGCAATTCCTATCCCTG 665  
 Db 514 TATTAATGCTATGCTGCTGCAATTTATGACCTAATTTGGCGCAATTCCTATCCCTG 455  
 Oy 666 ACTGGGAGTTCCTTAAGCTGCTGATGATGTTTTCGCAATTTGATGCAATTA 725  
 Db 454 ACTGGGAGTTCCTTAAGCTGCTGATGATGTTTTCGCAATTTGATGCAATTA 395  
 Oy 726 TGGCTATCATAAATTAATTCAGCTTAAGGAAACATCTTTCACGCCCTTTATAGTTGCT 785  
 Db 394 TGGCTATCATAAATTAATTCAGCTTAAGGAAACATCTTTCACGCCCTTTATAGTTGCT 335  
 Oy 786 GCAAGCAGCTTCTTAACCTGGGCTCATACCTGGAGACATCGTGGGAAAGATTAAG 845  
 Db 334 GCAAGCAGCTTCTTAACCTGGGCTCATACCTGGAGACATCGTGGGAAAGATTAAG 275  
 Oy 846 ACATGTAAGTCTTAAGGAGGCTGCGCAATTAATTAATTAATTAATTAATTAATTA 905  
 Db 274 ACATGTAAGTCTTAAGGAGGCTGCGCAATTAATTAATTAATTAATTAATTAATTA 215  
 Oy 906 AACCGGAATGAGATCTCATTTGAAAAAATATATGATGATGATGATGATGATGATGATG 965  
 Db 214 AACCGGAATGAGATCTCATTTGAAAAAATATATGATGATGATGATGATGATGATGATG 155  
 Oy 966 TAGGGGACCTTATTTATTTGCTTAACTGAATAGGAAATGATGATGATGATGATGATG 1025  
 Db 154 TAGGGGACCTTATTTATTTGCTTAACTGAATAGGAAATGATGATGATGATGATGATG 95  
 Oy 1026 TGAATGAGTATTTTTCCTTAAAGAGAAATGATGATGATGATGATGATGATGATGATG 1085  
 Db 94 TGAATGAGTATTTTTCCTTAAAGAGAAATGATGATGATGATGATGATGATGATGATG 35  
 Oy 1086 AAAAAAAAA 1094  
 Db 34 AAAAAAAAA 26

RESULT 5  
 AM812970/c 617 bp mRNA linear EST 17-MAY-2000  
 LOCUS AM812970  
 DEFINITION RC3-ST0186-230300-019-406 ST0186 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION AM812970  
 VERSION AM812970.1 GI:7905964  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 DIES Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brlones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalha, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunsfeld, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 JOURNAL MEDLINE  
 COMMENT 10737800  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP.

Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-RC3-ST0186-230  
 300-019-a066t3-2000-03-23&tl=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 32  
 High quality sequence stop: 617.

FEATURES  
 source location/Qualifiers  
 1. 617  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="ST0186"  
 /note="Organ: stomach; Vector: puc18; site\_1: SmaI;  
 site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

BASE COUNT 205 a 130 c 95 g 187 t  
 ORIGIN

Query Match 48.3%; Score 528; DB 9; Length 617;  
 Best Local Similarity 98.0%; Pred. No. 1e-102;  
 Matches 545; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

Oy 531 GAGAAAGAGTGATATTTGCTATGATGAGAGCTGCTGTTGTAATTAACGCTA 590  
 Db 617 GAGAAAGAGTGATATTTGCTATGATGAGAGCTGCTGTTGTAATTAACGCTA 558  
 Oy 591 TCCCTTGAATGCAATTTATTCGGCATGCTGCAATTTCAATTAATTAATTAATTA 650  
 Db 557 TCCCTTGAATGCAATTTATTCGGCATGCTGCAATTTCAATTAATTAATTAATTA 498  
 Oy 651 CAATTCATACCTGACTGGGAGTCTTAAGCTGCTGATGATGATGATGATGATGATG 710  
 Db 497 CAATTCATACCTGACTGGGAGTCTTAAGCTGCTGATGATGATGATGATGATGATG 438  
 Oy 711 TTGGATTCATTAATGCTATCATATAAATAATTAAGCTTAAGGAAACATCTTCAAC 770  
 Db 437 TTGGATTCATTAATGCTATCATATAAATAATTAAGCTTAAGGAAACATCTTCAAC 378  
 Oy 771 GCCTTAATGCTGCTCAGACAGCTTCTAAGCTGGGCTCATACCTGGAGAACATCTG 830  
 Db 377 GCCTTAATGCTGCTCAGACAGCTTCTAAGCTGGGCTCATACCTGGAGAACATCTG 318  
 Oy 831 GGGAAAGATTAAGCATGATGATGCTTAAGGAGGCTGGCATGAATTAACCTACTG 890  
 Db 317 GGGAAAGATTAAGCATGATGATGCTTAAGGAGGCTGGCATGAATTAACCTACTG 258  
 Oy 891 TTAGTGAGCAGCAAAACCGAATGAGATCCATGAAAAAATTAATGATTTGTAATG 950  
 Db 257 TTAGTGAGCAGCAAAACCGAATGAGATCCATGAAAAAATTAATGATTTGTAATG 198  
 Oy 951 TGAATTTTAAATTAAGGAGGACCTTATTTATTTGCTGTTAACTGAATAGGAAATG 1010  
 Db 197 TGAATTTTAAATTAAGGAGGACCTTATTTATTTGCTGTTAACTGAATAGGAAATG 138  
 Oy 1011 TACATACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1069  
 Db 137 TACATACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 78  
 Oy 1070 AAATGCTGTTTATTA 1085  
 Db 77 AAATGCTGTTTATTA 62

RESULT 6	530 bp	mrna	linear	EST_06-MAR-2002
BM826722				
LOCUS				
DEFINITION	BM826722			
ACCESSION	K-EST0098878	SI4K40ZS1	Homo sapiens	CDNA clone SI4K40ZS1-16-B01 5' /
VERSION	BM826722			
KEYWORDS	BM826722.1	GI:19183131		
SOURCE	EST.			
ORGANISM	Homo sapiens	(human)		
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
	1 (bases 1 to 530)			
	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,			
	Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and			
	Kim,I.S.			
TITLE	Z1C Frontier Korean EST Project 2001			
JOURNAL	Unpublished			
COMMENT	Contact: Kim YS			
	Genome Research Center			
	Korea Research Institute of Bioscience & Biotechnology			
	52 Eeundong, Yuseong-gu, Daejeon 305-333, South Korea			
	Tel: +82-42-860-4470			
	Fax: +82-42-860-4409			
	Email: yongsung@email.kribb.re.kr			
	Plate: 16 row: B column: 01			
FEATURES	High quality sequence stop: 530.			
source	Location/Qualifiers			
	1..530			

BASE COUNT	131 a	88 c	130 g	181 t
ORIGIN				

Query Match	46.38;	Score 507;	DB 12;	Length 530;
Best Local Similarity	99.88;	Pred. No. 3.1e-98;		
Matches 518;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1;

QY 292 TCAAGATTATTTC CCAAGTGATGAAGAAATGAGGGTCCCATTAAC TTGGGCGTCG 35

Db 12 TCAGATTATTTCCCAAGTGTGAAAGAAATGAGGGTTCACATTTGGGCTGCTG 71

QY 352 CTTGGTTTTTCTCCCTTGCTCGTCTGTGTGACTCAGGCTGGAATTTACTGGGTCA 411

Db 72 CTTGGTTTTTCTCCCTTGCTCGTCTGTGTGACTCAGGCTGGAATTTACTGGGTCA 131

QY 412 TCTGATTGACACCTTGCTGCTGGATGGGCAATTTTAATTCAGCTAATCTGGAGCTACT 471

Db 132 TCTGATTTGACACCTTGCTGCTGGATGGGCAATTTTAATTCAGCTAATCTGGAGCTAGT 191

QY 472 TGGATATCATCTGCATTTATGAGAGGAACAATTCATTGAGAGTACAAATGATGATGG 531

Db 192 TGGATATCATCTGCATTTATGAGAGGAACAATTCATTGAGAGTACAAATGATGATGG 251

QY 532 AGCAAGAGGTGATATTTCTGGCTATGATGAGAGGCTTGCTGTTGTATTT -ACGGCTA 590

Db 252 AGCAAGAGGTGATATTTCTGGCTATGATGAGAGGCTTGCTGTTGTATTTAATTAAGGCTA 311

QY 591 TCCCTTTGATTCGATATTTATCTGAGCATTTGGTGCATTTTCATAGACCTAATTTATGAGCG 650

Db 312 TCCCTTTGATTCGATATTTATCTGAGCATTTGGTGCATTTTCATAGACCTAATTTATGAGCG 371

QY 651 CAATTTCATACCTTCGACTGGGGAGTTGCTTTAGAGCTGGTGTATGATTTGTTTTCGCATTA 710

Db 372 CAATTTCATACCTTCGACTGGGGAGTTGCTTTAGAGCTGGTGTATGATTTGTTTTCGCATTA 431

QY 711 TTTGGATTCGCAATTTATGGCTATCATATAAATAATTAATTCAGGCTAAGGAAACATCTTTCAC 770

Db 432 TTTGGATTCGCAATTTATGGCTATCATATAAATAATTCAGGCTAAGGAAACATCTTTCAC 491

QY 771 GCCTATTAAGTTGTGCAGACAGCTTCTAATCTGGGCTC 809

Db 492 GCCTATTAAGTTGTGCAGACAGCTTCTAATCTGGGCTC 530

RESULT 7	691 bp	MRNA	linear	EST 12-JUN-2001
LOCUS	BG964290			
DEFINITION	602829090f1 NCI-CGAP-Co24	Mus musculus	cdna clone	IMAGE:4983526 5'
ACCESSION	BG964290			
VERSION	BG964290.1	GI:14351927		
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	1 (bases 1 to 691)			
TITLE	NH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .			
JOURNAL	National Institutes of Health. Mammalian Gene Collection (MGC)			
COMMENT	Unpublished			
	Contact: Robert Strausberg, Ph.D.			
	Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a>			
	Tissue Procurement: Jeffrey E. Green, M.D.			
	CDNA Library Preparation: Life Technologies, Inc.			
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)			
	DNA Sequencing by: Incyte Genomics, Inc.			
	DNA distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LNL at:			
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>			
	plate: LLMI0988 row: d column: 23			
	High quality sequence stop: 689.			
	location/Qualifiers			
	1..691			



Site\_2: Sali; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.6 kb. Constructed by Life  
 Technologies. Note: this is a NCI-CGAP Library."

BASE COUNT 191 a 124 c 160 g 216 t

Query Match 43.6%; Score 476.6; DB 12; Length 691;  
 Best Local Similarity 84.8%; Pred. No. 1.1e-91;  
 Matches 569; Conservative 0; Mismatches 99; Indels 3; Gaps 3;

QY 264 CCTGTGAGAAAGACACACACACAAATTCAGATTATTTCCCAAGTGAAGAAAA 323  
 Db 24 CTCTATTGAAACAATTAACACACATTTTCAAGATTGTTCCCAAGCAATGAACCA 83  
 QY 324 TGAAGGCTCCATTAACCTTGGCTGCTGCTGTTTCTCCCTGGCTGCTGCTG 383  
 Db 84 TGAAGGCTCCATTAAC- TTGGGATGCTGCTT- GATTGTTCTCCCTGGCTGCTGCTG 141  
 QY 384 TGACCTAGGCTGGAATTTACTGGGTTGATCTGATGACCACTTCTGCTGATGGGCA 443  
 Db 142 TGACCTAGGCTGGAATTTACTGGGTTGATCTGATGACCACTTCTGCTGATGGGCA 201  
 QY 444 TTTTAATTCAGCTTACTGAGAGTGTGGAATCATCTGATTTATGAGGAAACAGAT 503  
 Db 202 TTTTAATTCAGCTTACTGAGAGTGTGGAATCATCTGATTTATGAGGAAATAGAT 261  
 QY 504 TCATTGAGATACAGAAATGATGATTTGGAGCAAGAGGTGATTTCTGGCTATGGTGA 563  
 Db 262 TCATTGAGATATAGAAATGATGATTTGGAGCAAGAAAGATTTCTGGCTATGGTGA 321  
 QY 564 GAGCTTGGCTGTTGTAATTTAGCCTATCTTGTGATTTGCAATTTATCTGGCTATTG 623  
 Db 322 GAGCTTGGCTGTTGTAATTTAGCCTATCTTGTGATTTGCAATTTATCTGGCTATTG 381  
 QY 624 TGCAATTTCAATAGACCTAATTTATGGCGAATTTCCATACCTGAGCGGGAGTTCCTAG 683  
 Db 382 TGAATTTCAATAGACCTAATTTATGGCGAATTTCCATACCTGAGCGGGAGTTCCTAG 441  
 QY 684 GCTGCTGATGATTTTTCGCAATTTATTTGATTTGCAATTTATGCTATCAATTAATAA 743  
 Db 442 GCTGCTGATGATTTTTCGCAATTTATTTGATTTGCAATTTATGCTATCAATTAATAA 501  
 QY 744 TTCAGGCTTAAGAAACATCTTTCAACGCTTATTAAGTGTGAGACGAGCTTCACT 803  
 Db 502 TTCAGGCTTAAGAAACATCTTTCAACGCTTATTAAGTGTGAGACGAGCTTCACT 561  
 QY 804 GGGGCTTACCTGGAACACATCTGTTGGGGAAGATATAAGACATGTTGCTCTAATA 863  
 Db 562 GGGGCTTACCTGGAACACATCTGTTGGGGAAGATATAAGACATGTTGCTCTAATA 621  
 QY 864 AAGAGGCTGAGCAATACCTGTTAGTGTGAGACAGAAAAACCGAATGAGATCTCA 923  
 Db 622 AAGAGGCTGAGCAATACCTGTTAGTGTGAGACAGAAAAACCGAATGAGATCTCA 680  
 QY 924 TTGAAAAAAT 934  
 Db 681 CCTTTGAAAAAT 691

RESULT 8  
 AA552658/c 487 bp mRNA linear EST 05-SEP-1997  
 LOCUS n114b12.s1 NCI-CGAP\_C02 Homo sapiens cDNA clone IMAGE:1013471.3,  
 DEFINITION similar to SW:NTPR\_RAT P28573 SODIUM-DEPENDENT PROLINE TRANSPORTER.  
 ; mRNA sequence.  
 ACCESSION AA552658  
 VERSION AA552658.1 GI:2322912  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 487)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 JOURNAL Tumor Gene Index  
 COMMENT Unpublished

Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps@emall.nih.gov](mailto:cgaps@emall.nih.gov)  
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: Stratagene, Inc., David B. Kitzman,  
 Ph.D.

cDNA Library Arraying: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 found through the I.M.A.G.E. Consortium/LINL at:  
[www-bio.linnl.gov/bdrip/image/image.html](http://www-bio.linnl.gov/bdrip/image/image.html)  
 Insert Length: 1547 Std Error: 0.00  
 Seq primer: -40ml3 fwd. bp from Amersham  
 High quality sequence stop: 289.  
 Location/Qualifiers

#### FEATURES

##### SOURCE

1..487  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1013471"  
 /tissue\_type="tumor"  
 /lab\_host="SOLR (kanamycin resistant)"  
 /note="Organ: colon; Vector: Bluescript SK-; Site\_1: ECORI  
 /note="Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo  
 dt. Bulk colon villous adenoma. 5' adaptor sequence: 5'  
 GAATTCGGCAGCG 3' 3' adaptor sequence: 5'  
 CTCGAGTTTCTTTTCTTTTCTTTT 3' Average insert size: 1.1 kb."  
 BASE COUNT 158 a 99 c 72 g 158 t  
 ORIGIN

Query Match 43.2%; Score 472.6; DB 9; Length 487;  
 Best Local Similarity 98.2%; Pred. No. 7.4e-91;  
 Matches 478; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 599 ATTGCAATATTTATTCGTCATTTGTCATTTGATAGACCTAATTTAGCGCAATTTCA 658  
 Db 487 ATTGCAATATTTATTCGTCATTTGTCATTTGATAGACCTAATTTAGCGCAATTTCA 428  
 QY 659 TACCCTGACTGGGAGTCTCTTAGCGCTGATGATGTTCTTCTGCAATTTTGGATT 718  
 Db 427 TACCCTGACTGGGAGTCTCTTAGCGCTGATGATGTTCTTCTGCAATTTTGGATT 368  
 QY 719 CCATTTATGCTTCAATTAATAATTAATTCAGGCTAAGAGAAACATCTTCAACGCTTTA 778  
 Db 367 CCATTTATGCTTCAATTAATAATTAATTCAGGCTAAGAGAAACATCTTCAACGCTTTA 308  
 QY 779 AGTTGCTGAGACGAGCTTCACTAGCGGGTCCATACCTGGAACAACATCTGAGGGAAGA 838  
 Db 307 AGTTGCTGAGACGAGCTTCACTAGCGGGTCCATACCTGGAACAACATCTGAGGGAAGA 248  
 QY 839 TATTAAGACATGTAATCTTAATAAAGAGGCTGGCAGTAATACCTACTGTTAGTGGC 898  
 Db 247 TATTAAGACATGTAATCTTAATAAAGAGGCTGGCAGTAATACCTACTGTTAGTGGC 188  
 QY 899 AGCAGAAACCGGATTTGAGATCTCATTTGAAAAAATATATGATTTGATGATTTT 958  
 Db 187 AGCAGAAACCGGATTTGAGATCTCATTTGAAAAAATATATGATTTGATGATTTT 128  
 QY 959 TTTAGAAATAGGGGACCTTATTTATTTGTTGTTTCACTGATGAAATGTAATCACTACT 1018  
 Db 127 TTTAGAAATAGGGGACCTTATTTATTTGTTGTTTCACTGATGAAATGTAATCACTACT 68  
 QY 1019 ATGTTCAATGATAGGCTGATTTTCTCCATTTAAGCAGAAATGCAATATAAATGTGCT 1078  
 Db 67 ATGTTCAATGATAGGCTGATTTTCTCCATTTAAGCAGAAATGCAATATAAATGTGCT 8  
 QY 1079 TTTTATA 1085  
 I I I I I



```

OY 521 ATGATGATTGAGCAAGAGTATTTCTGGCGTATGATGAGAGACCTTCTGCTGTTGTA 580
DB 494 ATGATGATTGAGCAAGAGTATTTCTGGCGTATGATGAGAGACCTTCTGCTGTTGTA 435
OY 581 ATTAGCCCTATTCCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 640
DB 434 ATCACTCTATACCTGTTCTGCAATTTCTACTGTCGTCACGCGTGAATTTTCAACAGCT 375
OY 641 AATTATGCGCAATTTCCATACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 700
DB 374 AATTATGCGTATTTCCATACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 315
OY 701 TTTCGATTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 760
DB 314 TTTCGATTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 255
OY 761 ATCTTTCACGCGCTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 820
DB 254 ATCTTTCACGCGCTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 195
OY 821 CACATGCTGGGGGAGAAAGATATTAAGACATGCTGATCTTAAAGAGGCTGGCCATGAA 880
DB 194 AACATGCTGGAGAGAGATATTAAGACATGCTGATGATGATGATGATGATGATGATGAT 135
OY 881 ATACCTGCTGTTAGTGGAGCAAGAAACCGGATGATGATGATGATGATGATGATGATGAT 937
DB 134 ATACCTGCTGTTAGTGGAGCAAGAAACCGGATGATGATGATGATGATGATGATGATGAT 75
OY 938 TCAATGATTAAGTATTTTATTTTATGATGATGATGATGATGATGATGATGATGATGAT 997
DB 74 TTTCTATTAAGTATTTTATTTTATGATGATGATGATGATGATGATGATGATGATGAT 15
OY 998 GAATAGGAAATGT 1011
DB 14 AACTAGGGGAGTCT 1

RESULT 11
BX523178 ID BX523178 standard; RNA; EST; 567 BP.
AC BX523178;
SV BX523178.1
XX
DT 27-MAY-2003 (Rel. 75, Created)
DT 27-MAY-2003 (Rel. 75, Last updated, Version 1)
DE R2PD Mus musculus cDNA clone IMAGP998B202628 - IMAGE:1051315 5' EST.
XX EST: expressed sequence tag.
XX Mus musculus (house mouse)
XX Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
XX Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
XX [1]
XX 1-567
XX Heil O., Ebert L., Neubert P., Peters M., Radelof U., Schneider D.,
XX Korn B.;
XX Submitted (28-MAY-2003) to the EMBL/GenBank/DBJ databases.
XX R2PD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Im Neuenheimer
XX Feld 580, D-69120 Heidelberg, Germany
XX RZPD: IMAGP998B202628.
XX RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
XX Mouse Unigeneset - RZPD2 (RZPDLIB No.981)
XX http://www.rzpd.de/Clonedards/cgi-bin/showLib.pl.cgi?response71libNo=981
XX Contact: Ina Rolfs
XX RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
XX Heubnerweg 6, D-14059 Berlin, Germany
XX Tel: +49 30 32639 101

```

```

CC Fax: +49 30 32639 111
CC www.rzpd.de
CC This clone is available royalty-free from RZPD;
CC contact RZPD (clone@rzpd.de) for further information.
CC Seq primer: 17, primer sequence: TAATACGACTCCTATTAAGG
XX Key
XX Location/Qualifiers
FH 1. 567
FH /db_xref="taxon:10090"
FH /note="Tissue obtained from 8 week old mouse. Colon was
FH harvested 72 hours after irradiation with 1400 Gys. 1st
FH strand cDNA was primed with a Not I - oligo(dt) primer
FH (5'GTGTAAGATGCGAAGTGGAGCGCGCCGCTTTT)
FH TTTT)
FH Eco RI adaptors (AATTCGTGCAATG), digested with NotI and
FH cloned into the NotI and EcoRI sites of the p773D-PacI
FH vector. Library constructed by R. Barstead (Oklahoma
FH Medical Research Foundation).
FH /organism="Mus musculus"
FH /clone="IMAGP998B202628"
FH /clone_lib="Barstead irradiated colon MPL-RB7"
FH /dev_stage="8 weeks"
FH /lab_host="DH10B"
SQ
Sequence 567 BP; 158 A; 100 C; 140 G; 169 T; 0 other;
Query Match 37.9%; Score 415; DB 4; Length 567;
Best Local Similarity 86.0%; Pred. No. 1.7e-78;
Matches 460; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
OY 389 CAGCTGGAAATTTACTGGGCTCATGATGACCACTTGTGCTGATGGGCAATTTTA 448
DB 26 CAGCTGGAAATTTACTGGGCTCATGATGACCACTTGTGCTGATGGGCAATTTTA 85
OY 449 ATTGACGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 508
DB 86 ATTGACGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 145
OY 509 GAGGATACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 568
DB 146 GAGGATACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 205
OY 569 TGCTGGTGTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 628
DB 206 TGCTGGTGTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 265
OY 629 TTTCAATGACCTAATTTATGCGCAATTCATACCTGACCTGGGAGTTGCTGAGGCTG 688
DB 266 TTTCAATGACCTAATTTATGCGCAATTCATACCTGACCTGGGAGTTGCTGAGGCTG 325
OY 689 TGTATGATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 748
DB 326 TGTATGATGATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 385
OY 749 GCTAAAGGAAACATCTTTCAACGCGCTTATTAAGTGTGACAGACCAAGTTTCACTGGGCT 808
DB 386 GCTAAAGGAAACATCTTTCAACGCGCTTATTAAGTGTGACAGACCAAGTTTCACTGGGCT 445
OY 809 CCATACCTGGAACAACTGCGGGGGAAGATATTAAGACATGATGATGATGATGATGATGATGAT 868
DB 446 CCATACCTGGAACAACTGCGGGGGAAGATATTAAGACATGATGATGATGATGATGATGATGAT 505
OY 869 GCTGCGCATGAAATACCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 923
DB 506 ACTGACCAAGAAATACCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 560

```

```

RESULT 12
CB586665/c 582 bp mRNA linear EST 03-APR-2003
LOCUS AMGNNUC:SRC51-00008-G9-A src51 (10883) Rattus norvegicus cDNA clone
DEFINITION src51-00008-g9 5', mRNA sequence.

```

ACCESSION	CB586665	GI:29530706
VERSION	CB586665.1	
KEYWORDS	EST	
SOURCE	Rattus norvegicus (Norway rat)	
ORGANISM	Rattus norvegicus	
REFERENCE	1 (bases 1 to 582)	
AUTHORS	Angen Est Program	
TITLE	Angen Rat EST Program	
JOURNAL	unpublished	
COMMENT	Contact: Dan Fitzpatrick Angen, Inc One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA Tel: 805 447-4881 Plate: 00008 row: 9 column: 9	
FEATURES	Location/Qualifiers	
source	1..582	
	/organism="Rattus norvegicus"	
	/mol_type="mRNA"	
	/db_xref="taxon:10116"	
	/clone="sircs1-00008-g9"	
	/tissue_type="Cornea"	
	/clone_1ib="sircs1 (10883)"	
	/note="Vector: PCR4-TOPO; site_1: EcoRI; site_2: EcoRI; rat corneas"	
BASE COUNT	182 a 139 c 92 g 169 t	
ORIGIN		
Query Match	37.6%; Score 411.6; DB 14; Length 582;	
Best Local Similarity	82.8%; Pred. No. 8.9e-78;	
Matches	482; Conservative 0; Mismatches 99; Indels 1; Gaps 1	
QY	398 ATTATCTGGGTCATCTGATTTGACCCACTTCTGTCTGTGATGGGACATTTTAAATTCGACGT	457
DB	582 ATTACTGGGTTTCATTTGATTTGATTCATCTCTGTGCGATGGGGATCTTAATTCGTGCGC	523
QY	458 ATACCTGAGCTACTTGGAAATCATCTGGATTTATGAGGAGACAGATTCATTCAGATGATACA	517
DB	522 ATCTTGAATTCACAGAAATCATCTGATTTATGAGGAGAAATAGCTTCATTTGAATGATATA	463
QY	518 GAATGATGATTTGGAGCAAGAGAGGTGATATTTCTGCTATGTGTGAGAGACTTCCTGGTTT	577
DB	462 GAATGATGATTTGGAGCAAGAAAGATGCTATTTCTGCTTTGGTGGAGAGCTTGTGTGTTT	403
QY	578 GTATTTACGGCTATCTTTGATTTGATTCGCAATTTATCTGTCTGATTTGTCATTTTCATAGA	637
DB	402 GTCTATCACTCTATACCTGTTGTGTGCAATTCATGTCGTGACACGGTGAATTTTCACAGA	343
QY	638 CCTAATTTATGGCGCAATTCATTCATCCCTGACTGGGAGATTCCTTTAGCGTGGTGTATGAT	697
DB	342 CCTAATTTATGCTGATTTATTCATATCCCTGACTGGGAGATTCCTTCGGATGCTATGAT	283
QY	698 GTTTTCGCAATTTATTTGGATTCATTCATTTATGGCTTCATTAATAAATTAATTCAGGCTAAGA	757
DB	282 ATTTTCGCAATTTATTTGGATTCATTCATTTATGGCGATTCATTAATAAATTAATTCAGGCTAAGA	223
QY	758 AACATCTTTCAACGCTTATATAGTTGCTGCGAGACAGCTTCTACTGGGGCCATACCTG	817
DB	222 AACATCTTTCAACGCTTATATATATGCTGTGACCAAGCTTCTAATACTGGGGTCCATACCTG	166
QY	818 GAACCAACATTCGTTGGGGAAGAATATTAAGACATCTGATTTCTTAATAAAGAGGCTGGCAT	877
DB	162 GAACCAACATTCGTTGGGGAAGAATATTAAGACATCTGATTTCTTAATAAAGAGGCTGGCAC	103
QY	878 GAATACCTACTGTTAGTGGCAGCAGCAAAACCCGGAATGAGATTCATTAATAAATTAATA	933
DB	102 GAATACCTACTGTTAGTGGCAGCAGCAAAACCCGGAATGAGATTCATTAATAAATTAATA	43
QY	938 TGATTTG-TTAAATGATTTTATTTTGAATAGAGGAGACCTT 978	
DB	42 GAATTTCTCTAATATGATTTTGTAGAGTACGAGGGAATTTT 1	

[illegible]



Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=412-QV4-BN0090-040  
500-199-a06&tl=2000-05-04&tl=1)

Seq primer: puc 18 forward  
High quality sequence start: 21  
High quality sequence stop: 360.  
Location/Qualifiers

## FEATURES

source

1..360

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="BN0090"

/note="Organ: breast, normal; Vector: puc18; Site.1: SmaI;

Site.2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

## BASE COUNT

119 a 83 c 61 g 97 t

## ORIGIN

## Query Match

29.8%; Score 325.8; DB 10; Length 360;

Best Local Similarity 96.4%; Pred. No. 2,1e-59;

Matches 344; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

```
QY 475 AATCATCTGATTTATGAGGAGACAGATTCATTGAGATACAGAAATGATGGAGC 534
    |||||||
DB 360 AATCATCTGATTTATGAGGAGACAGATTCATTGAGATACAGAAATGATGGAGC 301
    |||||||
QY 535 AAAGAGTGATTTCTGCTATGCTGAGAGAGCTTGCTGTTGTAATTAAGCCATCT 594
    |||||||
DB 300 AAAGAGTGATTTCTGCTATGCTGAGAGAGCTTGCTGTTGTAATTAAGCCATCT 241
    |||||||
QY 595 TTGATTCGAATTTATCTGCTATGCTGATTCATTCATTAATTAATGAGCGCAT 654
    |||||||
DB 240 TTGATTCGAATTTATCTGCTATGCTGATTCATTCATTAATTAATGAGCGCAT 181
    |||||||
QY 655 TCCATACCTGCTGAGGAGTTGCTTTAGGCTGGTGTATGATGTTTCTGCAATTAT 714
    |||||||
DB 180 TCCATACCTGCTGAGGAGTTGCTTTAGGCTGGTGTATGATGTTTCTGCAATTAT 121
    |||||||
QY 715 GATTCGAATTTATGCTATTCATTAATAATTAATTAATTAATTAATTAATTAAT 774
    |||||||
DB 120 GATTCGAATTTATGCTATTCATTAATAATTAATTAATTAATTAATTAATTAAT 61
    |||||||
QY 775 TATTAAGTTCGACAGCAGCTTCTAAGTGGGCTCATATACCTGGAACAACATCG 831
    |||||||
DB 60 TATTAAGTTCGACAGCAGCTTCTAAGTGGGCTCATATACCTGGAACAACATCG 5
    |||||||
```

Search completed: August 27, 2003, 12:07:17  
Job time : 2784 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 17:32:44 ; Search time 29 Seconds

(without alignments)  
315.143 Million cell updates/sec

Title: US-09-940-227-83

Perfect score: 1198  
Sequence: 1 SEETITTTIIDLPRVKKMKM.....VPRKEAGHEIPTVSGSRKPE 216

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	504.5	42.1	799 2	US-08-700-013B-27 Sequence 27, Appl
2	503.5	42.0	797 3	US-09-182-728A-2 Sequence 2, Appl
3	503.5	42.0	797 4	US-09-191-468-120 Sequence 120, App
4	503.5	42.0	797 4	US-09-191-468-122 Sequence 122, App
5	503.5	42.0	797 4	US-09-795-232-2 Sequence 2, Appl
6	502.5	41.9	280 2	US-08-700-013B-17 Sequence 17, Appl
7	502.5	41.9	797 2	US-08-700-013B-19 Sequence 19, Appl
8	502.5	41.9	797 4	US-09-191-468-124 Sequence 124, App
9	501.5	41.9	280 2	US-08-700-013B-15 Sequence 15, Appl
10	501.5	41.9	797 2	US-08-700-013B-21 Sequence 21, Appl
11	462.5	38.6	194 2	US-08-700-013B-25 Sequence 25, Appl
12	447.5	37.4	194 2	US-08-700-013B-23 Sequence 23, Appl
13	359	30.0	667 1	US-07-879-617A-8 Sequence 8, Appl
14	359	30.0	667 1	US-08-753-985-8 Sequence 8, Appl
15	350.5	29.3	150 4	US-09-191-468-88 Sequence 92, Appl
16	350.5	29.3	150 4	US-09-191-468-92 Sequence 92, Appl
17	350.5	29.3	150 4	US-09-191-468-94 Sequence 94, Appl
18	350.5	29.3	150 4	US-09-191-468-96 Sequence 96, Appl
19	350.5	29.3	150 4	US-09-191-468-98 Sequence 98, Appl
20	350.5	29.3	150 4	US-09-191-468-100 Sequence 100, App
21	350.5	29.3	150 4	US-09-191-468-102 Sequence 102, App
22	345.5	28.8	687 3	US-08-834-467-2 Sequence 2, Appl
23	345.5	28.8	687 3	US-09-396-177-2 Sequence 2, Appl
24	344.5	28.8	150 4	US-09-191-468-90 Sequence 90, Appl
25	344.5	28.8	638 1	US-08-295-814E-13 Sequence 13, Appl
26	344.5	28.8	638 1	US-08-240-783B-2 Sequence 2, Appl
27	344.5	28.8	638 3	US-09-084-813-2 Sequence 2, Appl

28	344.5	28.8	638 3	US-09-343-361-13 Sequence 13, Appl
29	344.5	28.8	638 5	PCT-US92-09662-2 Sequence 2, Appl
30	341	28.5	635 1	US-07-879-617A-9 Sequence 9, Appl
31	341	28.5	635 1	US-08-753-985-9 Sequence 9, Appl
32	325	27.1	616 1	US-07-879-617A-13 Sequence 13, Appl
33	325	27.1	616 1	US-08-753-985-13 Sequence 13, Appl
34	319	26.6	620 1	US-08-301-722A-2 Sequence 2, Appl
35	316	26.4	619 1	US-07-762-132A-2 Sequence 2, Appl
36	316	26.4	619 1	US-08-301-722A-2 Sequence 2, Appl
37	308.5	25.8	599 1	US-07-879-617A-10 Sequence 10, Appl
38	308.5	25.8	599 1	US-08-295-814E-11 Sequence 11, Appl
39	308.5	25.8	599 1	US-08-240-783B-4 Sequence 4, Appl
40	308.5	25.8	599 3	US-08-753-985-10 Sequence 10, Appl
41	308.5	25.8	599 3	US-09-084-813-4 Sequence 4, Appl
42	308.5	25.8	599 3	US-09-343-361-11 Sequence 11, Appl
43	308.5	25.8	599 5	PCT-US92-09662-4 Sequence 4, Appl
44	299	25.0	617 1	US-07-879-617A-11 Sequence 11, Appl
45	299	25.0	617 1	US-08-301-722A-3 Sequence 3, Appl

#### ALIGNMENTS

RESULT 1  
US-08-700-013B-27  
Sequence 27, Application US/08700013B  
Patent No. 5919653  
GENERAL INFORMATION:  
APPLICANT: Albert, Vivian R.  
APPLICANT: Kowalski, Leslie R. Z.  
APPLICANT: Borden, Laurence A.  
APPLICANT: McElvey, Jeffrey F.  
TITLE OF INVENTION: Human Glycine Transporter  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Dechert Price & Rhoads  
STREET: 997 Lenox Drive, Building 3, Suite 210  
CITY: Lawrenceville  
STATE: NJ  
COUNTRY: USA  
ZIP: 08543  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/700,013B  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bloom, Allen  
REGISTRATION NUMBER: 29,135  
REFERENCE/DOCKET NUMBER: 317743-108  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-520-3214  
TELEFAX: 609-520-3259  
TELEX:  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 799 amino acids  
TYPE: amino acids  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-700-013B-27  
Query Match 42.1%; Score 504.5; DB 2; Length 799;  
Best Local Similarity 46.4%; Pred. No. 1;le-46;  
Matches 91; Conservative 33; Mismatches 65; Indels 3; Gaps 1;

[illegible]



```

? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Bloom, Allen
? REGISTRATION NUMBER: 29, 135
? REFERENCE/DOCKET NUMBER: 317743-108
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 609-520-3214
? TELEFAX: 609-520-3259
? TELEX:
? INFORMATION FOR SEQ ID NO: 17:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 280 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
?
US-08-700-0138-17

Query Match 41.9%; Score 502.5; DB 2; Length 280;
Best Local Similarity 46.4%; Pred. No. 5e-47;
Matches 91; Conservative 33; Mismatches 69; Indels 3; Gaps 1

```

```

Db      71  EIVIVISISDEPKYLRTAKRPVETLGCCICEFFMGPMITGGIYMQVLVDYLAASVALVI 130
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::||||:::
QY      63  AALLETVOGIWYGSNGRPIEDTETMMIAKKNIMFWLMMRACWPFYITILLAIPIMSLOY 122
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::||||:::
Db      131  IAIPELVGISTYVGLQRFCEDEIMMISFOPNIF--WKVCMAFVPTPTITITFCISFYOW 187
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::||||:::
QY      123  HRPNTGAIPYDPMGVALGCMCIYVCIIPIFIMAIIRIKIIOAKGNIFORLISCCRPASNGP 182
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::||||:::
Db      188  EPMFTGSGYRPDMSVGLGMLLACQSVIWMIDIMVEYKMHLLAPGRFIERLTVCSPPDMGP 247
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::||||:::
QY      183  YLEQHRGERRYKDMVYP 198
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::
Db      248  FLAQRHGERRYKRMNDP 263
      |||:::||||:::||||:::||||:::||||:::||||:::||||:::

RESULT 7
US-08-700-013B-19
: Sequence 19, Application US/08700013B
: Patent No. 5010552

```

APPLICANT: Albert, Vivian R.  
APPLICANT: Kowalski, Leslie R.Z.  
APPLICANT: Borden, Laurence A.  
APPLICANT: McKelvy, Jeffrey F.  
TITLE OF INVENTION: Human Glycine Transporter  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert Price & Rhoads  
STREET: 997 Lenox Drive, Building 3, Suite 210  
CITY: Lawrenceville.  
STATE: NJ  
COUNTRY: USA  
ZIP: 08543  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Fastseq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/700,013B  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD M. HARRIS  
FIRM: HARRIS, HARRIS & HARRIS, P.C.  
ADDRESS: 1000 Peachtree Street, N.E.  
CITY: Atlanta, Georgia 30309  
PHONE: 404/525-8800  
FAX: 404/525-8801  
TELETYPE: 404/525-8802  
TELEFAX: 404/525-8803  
E-MAIL: RICHARD@HARRISHARRIS.COM  
HARRIS, HARRIS & HARRIS, P.C.  
ATTORNEY/AGENT INFORMATION:

```

1 REGISTRATION NUMBER: 29,135
2 REFERENCE/DOCKET NUMBER: 317743-108
3 TELECOMMUNICATION INFORMATION:
4 TELEPHONE: 609-520-3214
5 TELEFAX: 609-520-3259
6
7 TELEX:
8 INFORMATION FOR SEQ. ID NO.: 19:
9 SEQUENCE CHARACTERISTICS:
10 LENGTH: 797 amino acids
11 TYPE: amino acid
12 STRANDEDNESS: single
13 TOPOLOGY: linear
14
15 NS-08-700-013b-19

```

Query Match	41.9%	Score 502.5:	DB 2,	Length 797;
Best Local Similarity	46.4%:	Pred. No. 1,8e-46;		
Matches 91;	Conservative 33;	Mismatches 69;	Indels 3;	Gaps 1;

QY	ETTTTTTIOLEPVRKKMKKVPRTTACCCVLTFLGIVCTQAGIYFWHLIHEFCAGMILL 62
	:  :
Db	584 ETTVTSTISDEFFKYLRTFKRPVTTGCCCTCFIIMGFPMTTGGIYMFQVLDTPAASVLYI 643
	:  :
QY	63 AATLELVGIIMVYGNRFIEDTMMIGAKRWIEMWBRACVFTPTILLIAIFWLSVQF 122
	:  :
Db	644 IATFELVGISSVYGLQRFCEDEIMMIGQPNIF--MKVCWAFVPTLIFELCFSEFYQ 700
	:  :
QY	123 HRPNTGALPYPPDMGVALGWMCTVFCITLWIPMAITKIITQAKGNIFORLISCRASWMP 182
	:  :
Db	701 EPMTYSYSTRPNKSNMVLGWLMLACSVMIPIWFIKMLAARGREIRLKIYCSQPQMPG 760
	:  :
QY	183 YLEOHGGERKMDVVP 198
	::  :
Db	761 FTLAÖHGERKKNMIDP 776

```

RESULT 8
US-09-191-468-124
Sequence 124, Application US/09191468A
Patent No. 6416975
GENERAL INFORMATION:
APPLICANT: Gallagher, Michael J.
APPLICANT: Burgess, Loyd R.
APPLICANT: Brunden, Kurt R.
TITLE OF INVENTION: Human Glycine Transporter Type 2
FILE REFERENCE: 12311US01
CURRENT APPLICATION NUMBER: US/09/191,468A
CURRENT FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 124
SOFTWARE: Patentl In Ver. 2.0
SEQ ID NO 124
LENGTH: 797
TYPE: PRT
ORGANISM: Human
US-09-191-468-124

```

[illegible]

Db 761 FLAQHGRERYKNMIDP 776

RESULT 9  
 US-08-700-013B-15  
 Sequence 15, Application US/08700013B  
 Patent No. 5916653  
 GENERAL INFORMATION:  
 APPLICANT: Alibert, Vivian R.  
 APPLICANT: Kowalski, Leslie R.Z.  
 APPLICANT: Borden, Laurence A.  
 APPLICANT: McKeiliv, Jeffrey F.  
 TITLE OF INVENTION: Human Glycine Transporter  
 NUMBER OF SEQUENCES: 41  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Dechart Price & Rhoads  
 STREET: 997 Lenox Drive, Building 3, Suite 210  
 CITY: Lawrenceville  
 STATE: NJ  
 COUNTRY: USA  
 ZIP: 08543  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSD for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/700.013B  
 FILING DATE:  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bloom, Allen  
 REGISTRATION NUMBER: 29,135  
 REFERENCE/DOCKET NUMBER: 317743-108  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 609-520-3214  
 TELEFAX: 609-520-3259  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 280 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-700-013B-15

Query Match	Similarity	41.9%	Score	501.5:	DB 2:	length	280;	
Best Local	Similarity	46.4%;	Pred	No. 6.4e-47:				
Matches	Conservative	91;	Mismatches	33;	Indels	3;	Gaps	1
QY	3	ETITTTIODELPKVKMKRNVPTTLCCLVLFLGLGVCTVQAGIYWHLIDHFCAGGILLI	62					
Db	71	ETIVASIDSEFPKVFRTKHVPFTLLCCGCCCFEIMGPMPTGGIYMFOLVDYAAASYALVI	130					
QY	63	AALIELLVGIWYGNNRFIEDENMIGAKRMTEIMWRACAFVITPILIIAIFIMSLOVE	122					
Db	131	IAIEFLVGISYYGIORFCEDEEMMGHPDNFE---WKVCMAFLVPTLITFLFCSFYOW	187					
QY	123	HRPNGAIPDPDWGVALWCMIWFIPIWIPIAIKIIOAKNIPORLIISCPAPSNMGP	182					
Db	168	EPMYTGVSRYRPRMSAVLGMWLACGSVMIPIMEIVIKMLDPAGRFERLIKLVCSPPQDMGP	247					
QY	183	VLEOHRGERERYKMAYVP	198					
Db	248	FLAOHRGERERYKNMIDP	263					
RESULT 10								
US-08-700-013B-21								
: sequence 21, Application US/08700013B								

```
Patent No. 5919653
GENERAL INFORMATION:
APPLICANT: Albert, Vivian R.
APPLICANT: Kowalski, Leslie R.Z.
APPLICANT: Borden, Laurence A.
APPLICANT: McKelvy, Jeffrey F.
TITLE OF INVENTION: Human Glycine Transporter
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
CITY: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,013B
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29,135
REFERENCE/DOCKET NUMBER: 317743-108
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX:
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 797 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-700-013B-21

Query Match          41.9%; Score 501.5; DB 2; Length 797;
Best Local Similarity 46.4%; Pred. No. 2.3e-46;
Matches 91; Conservative 33; Mismatches 69; Indels 3; Gaps 1;

QY 3 ETITTTODLPKVKMKRWRYPTLGCCLVLEGLVCVTOAGIYVWHLIDHFCAGMGLI 62
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 584 ETIVTISIDEPKYLRTKHPVFLGCGVCFIMGFPMITGGIGIMFQLDVTVYASVAVI 643
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 63 AAILELVGIWIVGNGFIEDTEMGAKRIFWLMRACWFVITPILLIAIFWISLVQF 122
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 644 IAIELVGISIVYGLQFCEDEIMMIGFQPNIF--WKVCWAFVTPILLIFILCFSEYQW 700
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 123 HRPVYGAIPYDVGWALGMCIVFCIIMIPIMAIKIKIOAKGNIFORLISCCRPASWMP 182
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 701 EPMYGSYRYPNMSVGLWMLACSVIMIMVIMIKHMLRGRFIERLKLVCSPQPDWGP 760
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 183 YLEOHGREGKRDWVP 198
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 761 FLAQRGERYKKNMIDP 776
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 11
US-08-700-013B-25
Sequence 25, Application US/08700013B
Patent No. 5919653
GENERAL INFORMATION:
APPLICANT: Albert, Vivian R.
APPLICANT: Kowalski, Leslie R.Z.
APPLICANT: Borden, Laurence A.
APPLICANT: McKelvy, Jeffrey F.
```

```
TITLE OF INVENTION: Human Glycine Transporter
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
CITY: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,013B
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29,135
REFERENCE/DOCKET NUMBER: 317743-108
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX:
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-700-013B-25

Query Match          38.6%; Score 462.5; DB 2; Length 194;
Best Local Similarity 47.1%; Pred. No. 7.5e-43;
Matches 82; Conservative 30; Mismatches 59; Indels 3; Gaps 1;

QY 25 TLGCCCLVLEGLVCVTOAGIYVWHLIDHFCAGMGLIAAILVGIWIVGNGFIEDTE 84
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 3 TLGCCCLVLEGLVCVTOAGIYVWHLIDHFCAGMGLIAAILVGIWIVGNGFIEDTE 62
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 85 EMHIGKRNIFWLMRACWFVITPILLIAIFWISLVQFHRPNYGAIPYDVGWALGMCW 144
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 63 EMHIGKRNIFWLMRACWFVITPILLIAIFWISLVQFHRPNYGAIPYDVGWALGMCW 119
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 145 VFCIINIPIMAIKIKIOAKGNIFORLISCCRPASWMPYLEOHGREGKRDWVP 198
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 120 ACSVIMIPIMSVYIKHMLRGRFIERLKLVCSPQPDWGPFLAQRGERYKKNMIDP 173
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 12
US-08-700-013B-23
Sequence 23, Application US/08700013B
Patent No. 5919653
GENERAL INFORMATION:
APPLICANT: Albert, Vivian R.
APPLICANT: Kowalski, Leslie R.Z.
APPLICANT: Borden, Laurence A.
APPLICANT: McKelvy, Jeffrey F.
TITLE OF INVENTION: Human Glycine Transporter
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
CITY: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:
```

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/700,013B  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bloom, Allen  
REGISTRATION NUMBER: 29,135  
REFERENCE/DOCKET NUMBER: 317743-108  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-520-3214  
TELEFAX: 609-520-3259  
TELEX:  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 194 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-700-013B-23

Query Match 37.4%; Score 447.5; DB 2; Length 194;  
Best Local Similarity 46.0%; Pred. No. 3.3e-41;

Matches 80; Conservative 30; Mismatches 61; Indels 3; Gaps 1;

QY 25 TLGCLVLFLLGLVCTOAGIYVHLIDHCAGMGLIAILVGLVIMVYGNRIEDT 84  
DB 3 TLGCIYCFIFMGFMITGGIYMFQLDVTAASAYALVILAFELVGISYVGLQRCEDI 62  
QY 85 EMTGARWTFWLRACWVPIPIILAFIMSLVOFHPRNGAIPYPMGVALGCM 144  
DB 63 EMTGAFQNTF--WKQCAVFTPTITFLICFSFYMEEMTYSYPRMSVGLGLML 119  
QY 145 VECIWIPIMAIKIIOAKNIFORLISCPASNMGPYIEQHRGEYKMKVYP 198  
DB 120 ACSYIPIPIVVMKHLAPGRFIRLKLVCSPQPMGPFLAHRGERYKMWIDE 173

RESULT 13  
US-07-879-617A-8  
Sequence 8, Application US/07879617A  
Patent No. 5580775  
GENERAL INFORMATION:  
APPLICANT: Fremau Jr., Robert T.  
APPLICANT: Caron, Marc G.  
APPLICANT: Blakey, Randy D.  
TITLE OF INVENTION: A High Affinity L-Proline Transporter  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: U.S.  
ZIP: 30309  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/879,617A  
FILING DATE: 19920501  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Padst, Patricia L.  
REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: EMU109  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6508  
TELEFAX: 404-815-6555  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 667 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Rattus  
TISSUE TYPE: Brain  
IMMEDIATE SOURCE:  
LIBRARY: rat forebrain cDNA library  
CLONE: rTB2-2-20  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 76..95  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 103..127  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 147..167  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 247..266  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 272..294  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 321..341  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 353..375  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 404..427  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 454..473  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 487..509  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 530..549  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 567..589  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 44..45  
OTHER INFORMATION: /note= "protein kinase C phosphorylation site"  
OTHER INFORMATION:

RESULT 14  
 US-08-753-985-8  
 Sequence 8, Application US/08733985  
 Patent No. 5753788  
 GENERAL INFORMATION:  
 APPLICANT: Fremneau Jr., Robert T.  
 APPLICANT: Marc G.  
 APPLICANT: Blakely, Randy D.  
 TITLE OF INVENTION: A High Affinity L-Proline Transporter  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Kilpatrick & Cody  
 STREET: 1100 Peachtree Street, Suite 2800  
 CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: U.S.  
 ZIP: 30309  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/753,985  
 FILING DATE: 03-DEC-1996  
 CLASSIFICATION: 536

1	PRIOR APPLICATION DATA:
2	APPLICATION NUMBER: US 07/879617
3	FLING DATE: 01-MAY-1992
4	ATTORNEY/AGENT INFORMATION:
5	NAME: Pabst, Patricia L.
6	REGISTRATION NUMBER: 31,284
7	REFERENCE/DOCKET NUMBER: EMU109
8	TELECOMMUNICATION INFORMATION:
9	TELEPHONE: 404-815-6508
10	TELEFAX: 404-815-6555
11	INFORMATION FOR SEQ ID NO: 8:
12	SEQUENCE CHARACTERISTICS:
13	LENGTH: 667 amino acids
14	TYPE: amino acid
15	STRANDEDNESS: single
16	TOPOLOGY: linear
17	MOLECULE TYPE: protein
18	HYPOTHETICAL: NO
19	ANTI-SENSE: NO
20	FRAGMENT TYPE: N-terminal
21	ORIGINAL SOURCE:
22	ORGANISM: Rattus
23	TISSUE TYPE: Brain
24	IMMEDIATE SOURCE:
25	LIBRARY: rat forebrain cDNA library
26	CLONE: RTB2-2-20
27	FEATURE:
28	NAME/KEY: Domain
29	LOCATION: 76..95
30	OTHER INFORMATION: /note= "Membrane-spanning domain"
31	FEATURE:
32	NAME/KEY: Domain
33	LOCATION: 103..127
34	OTHER INFORMATION: /note= "Membrane-spanning domain"
35	FEATURE:
36	NAME/KEY: Domain
37	LOCATION: 147..167
38	OTHER INFORMATION: /note= "Membrane-spanning domain"
39	FEATURE:
40	NAME/KEY: Domain
41	LOCATION: 247..266
42	OTHER INFORMATION: /note= "Membrane-spanning domain"
43	FEATURE:
44	NAME/KEY: Domain
45	LOCATION: 272..294
46	OTHER INFORMATION: /note= "Membrane-spanning domain"
47	FEATURE:
48	NAME/KEY: Domain
49	LOCATION: 321..341
50	OTHER INFORMATION: /note= "Membrane-spanning domain"
51	FEATURE:
52	NAME/KEY: Domain
53	LOCATION: 353..375
54	OTHER INFORMATION: /note= "Membrane-spanning domain"
55	FEATURE:
56	NAME/KEY: Domain
57	LOCATION: 404..427
58	OTHER INFORMATION: /note= "Membrane-spanning domain"
59	FEATURE:
60	NAME/KEY: Domain
61	LOCATION: 454..473
62	OTHER INFORMATION: /note= "Membrane-spanning domain"
63	FEATURE:
64	NAME/KEY: Domain
65	LOCATION: 487..509
66	OTHER INFORMATION: /note= "Membrane-spanning domain"
67	FEATURE:
68	NAME/KEY: Domain
69	LOCATION: 530..549
70	OTHER INFORMATION: /note= "Membrane-spanning domain"
71	FEATURE:
72	NAME/KEY: Domain
73	LOCATION: 567..589

OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 44..45  
OTHER INFORMATION: /note= "protein kinase C phosphorylation site"  
OTHER INFORMATION: phosphorylation site"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 71..72  
OTHER INFORMATION: /note= "cAMP-dependent protein kinase phosphorylation site"  
OTHER INFORMATION: kinase phosphorylation site"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 269..270  
OTHER INFORMATION: /note= "protein kinase C phosphorylation site"  
OTHER INFORMATION: phosphorylation site"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 630..631  
OTHER INFORMATION: /note= "protein kinase C phosphorylation site"  
OTHER INFORMATION: phosphorylation site"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 106..127  
OTHER INFORMATION: /note= "leucine zipper motif"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 212..213  
OTHER INFORMATION: /note= "N-linked glycosylation site"  
OTHER INFORMATION: site"  
US-08-753-985-8

Query Match 30.0%; Score 359; DB 1; Length 667;  
Best Local Similarity 30.4%; Pred. No. 6,9e-31;  
Matches 65; Conservative 55; Mismatches 78; Indels 16; Gaps 2;

QY 3 ETITTTTODLPKMKMRPITIGCCLVLFILGLVCTQAGIYVWHLIDHFCAGWILL 62  
DB 435 ETIVAVTDEFPYLRPRKAVFSGLICVAMTLMGLITTDGKMTVLDDYSASFGLMV 494  
QY 63 AAILELGIWYIGNRFIEDTEMIGAKRWIFWLMWRACFVITPILIAIFIMSLYOF 122  
DB 495 VVITTCCLAVTRYGIGRFRCDIHMLGFKP---GLYFRACWLFISPATLLALLVYSIVKY 551  
QY 123 HRPNTGAIPYDWCVALGCMIVFCIIMIPIMAIITIKIIIOAKGNIFORLISCCRPASNMGP 182  
DB 552 QPSEYGSYRFPAMALLGILMGLSLMTIPAGMLVAVIRGSLMERLQASRPALIDMP 611  
QY 183 VLEHGRERYKDMVVPKKEAGHEIPTVSGSRPE 216  
DB 612 SLEENR-----TGMYVATLAGSQSPK 632

RESULT 15  
US-09-191-468-88  
; Sequence 88, Application US/09191468A  
; Patent No. 6416975  
; GENERAL INFORMATION:  
; APPLICANT: Gallagher, Michael J.  
; APPLICANT: Burdgen, Loyd R.  
; TITLE OF INVENTION: Human Glycine Transporter Type 2  
; FILE REFERENCE: 12311US01  
; CURRENT APPLICATION NUMBER: US/09/191,468A  
; NUMBER OF SEQ ID NOS: 124  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 88  
; LENGTH: 150  
; TYPE: PRT  
; ORGANISM: Human  
US-09-191-468-88

Query Match 29.3%; Score 350.5; DB 4; Length 150;  
Best Local Similarity 48.5%; Pred. No. 9,7e-31;  
Matches 64; Conservative 20; Mismatches 45; Indels 3; Gaps 1;

QY 67 ELVGIITWYIGNRFIEDTEMIGAKRWIFWLMWRACFVITPILIAIFIMSLYOFHNP 126  
DB 1 ELVGISYVIGLORFCDIEDMIGFQPNIE---WKYCMAFVFPILITFLICSFYQWEDPT 57  
QY 127 YGAIIPPYDWCVALGCMIVFCIIMIPIMAIITIKIIIOAKGNIFORLISCCRPASNMGPYLEQ 186  
DB 58 YGSYRYPNMGSMVGLMILACSIVIMIPIMFVIRKMLHAPGRTLRKLVCSFQPDWGPFLAQ 117  
QY 187 HGRERYKDMVVP 198  
DB 118 HGRERYKMDIP 129

Search completed: August 27, 2003, 17:37:57  
Job time : 31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 17:36:34 ; Search time 52 Seconds

(Without alignments)  
568.171 Million cell updates/sec

Title: US-09-940-227-83

Perfect score: 1198  
Sequence: 1 SEETITTTTQDLFPKVKMKM.....VPKKEAGHEIPVSGSRKPE 216

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 510680 seqs, 136781880 residues

Total number of hits satisfying chosen parameters: 510680

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubppa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubppa/PCCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubppa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	503.5	42.0	797	9	US-09-795-232-2
2	361	30.1	636	12	US-10-289-161A-3
3	344.5	28.8	633	12	US-10-205-219-30
4	344.5	28.8	638	12	US-09-845-908-13
5	319	26.6	620	12	US-10-241-220-109
6	308.5	25.8	599	12	US-09-845-908-11
7	296	24.7	599	12	US-10-345-680-5
8	295.5	24.7	624	9	US-09-795-693-24
9	295.5	24.7	624	15	US-10-156-238-24
10	295.5	24.7	624	15	US-10-199-485-24
11	287.5	24.0	727	10	US-09-815-923-10
12	286.5	23.9	627	12	US-09-845-908-4
13	283.5	23.7	602	12	US-09-845-908-2
14	281.5	23.5	587	10	US-09-815-923-4
15	280	23.4	622	9	US-09-843-598-11

16	279.5	23.3	246	12	US-09-845-908-8	Sequence 8, Appl1
17	277.5	23.2	632	12	US-09-845-908-12	Sequence 10, Appl1
18	273	22.8	614	12	US-09-845-908-10	Sequence 12, Appl1
19	264	22.0	614	11	US-09-919-039-378	Sequence 378, Appl
20	255.5	21.3	621	12	US-09-845-908-6	Sequence 6, Appl1
21	251.5	21.0	635	12	US-10-205-194-33	Sequence 33, Appl1
22	248	20.7	633	12	US-10-298-974-2	Sequence 2, Appl1
23	248	20.7	633	12	US-09-991-458-2	Sequence 2, Appl1
24	246	20.5	597	10	US-09-815-923-14	Sequence 14, Appl1
25	243.5	20.3	630	9	US-09-843-598-10	Sequence 10, Appl1
26	242.5	20.2	730	9	US-09-741-149-2	Sequence 2, Appl1
27	242.5	20.2	730	9	US-09-795-693-5	Sequence 5, Appl1
28	242.5	20.2	730	12	US-10-385-614-2	Sequence 2, Appl1
29	242.5	20.2	730	15	US-10-156-239-5	Sequence 5, Appl1
30	242.5	20.2	730	15	US-10-199-485-5	Sequence 5, Appl1
31	235.5	19.7	729	9	US-09-741-149-4	Sequence 4, Appl1
32	235.5	19.7	729	12	US-10-385-614-4	Sequence 4, Appl1
33	222.5	18.6	610	10	US-09-861-846-2	Sequence 2, Appl1
34	222	18.5	676	10	US-09-815-923-12	Sequence 12, Appl1
35	210	17.5	599	10	US-09-861-846-4	Sequence 4, Appl1
36	209.5	17.5	727	9	US-09-923-444A-2	Sequence 2, Appl1
37	209.5	17.5	556	10	US-09-815-923-6	Sequence 6, Appl1
38	208	17.4	671	9	US-09-843-598-5	Sequence 5, Appl1
39	200	16.7	671	9	US-09-843-598-7	Sequence 7, Appl1
40	131	10.9	196	11	US-09-989-442-87	Sequence 87, Appl1
41	104	8.7	454	15	US-10-156-761-12378	Sequence 12378, A
42	100.5	8.4	579	10	US-09-738-626-4648	Sequence 1066, Ap
43	92	7.7	188	12	US-10-017-161-1046	Sequence 6706, Ap
44	89	7.4	463	10	US-09-738-626-6706	Sequence 56, Appl
45	88.5	7.4	501	10	US-09-934-866-56	

## ALIGNMENTS

RESULT 1	US-09-795-232-2	Application US/09795232
Sequence 2, Appl1	US-09-795-232-2	Patent No. US20010012627A1
GENERAL INFORMATION:		
APPLICANT:	Anthony M. Brown	
APPLICANT:	Conrad Gerald Chapman	
APPLICANT:	Israel Simon Gloger	
APPLICANT:	Joanne Rachel Evans	
APPLICANT:	William Cairns	
TITLE OF INVENTION:	NOVEL COMPOUNDS	
FILE REFERENCE:	GP-30176-D1	
CURRENT APPLICATION NUMBER:	US/09/795,232	
PRIOR FILING DATE:	2001-02-28	
PRIOR APPLICATION NUMBER:	09/182,728	
PRIOR FILING DATE:	1998-10-29	
PRIOR APPLICATION NUMBER:	9818890.7	
NUMBER OF SEQ ID NOS:	6	
SOFTWARE:	FastSeq for Windows Version 3.0	
SEQ ID NO 2		
LENGTH: 797		
TYPE: PRT		
ORGANISM: HOMO SAPIENS		
Query Match	42.08: Score 503.5; DB 9; Length 797;	
Best Local Similarity	46.48: Pred. No. 1.9e-42;	
Matches	91: Conservative 33; Mismatches 69; Indels 3; Gaps 1;	
QY	3 ETITTTTQDLFPKVKMKRVPITIGCLVFLGLVCVTDAGIYWHLDHFCAGGII 62	
DB	584 ETITTSISDEPKLRTKRFVFTLGCCICFCFIMGFPMTGGIYMQVDVTAASALVI 643	
QY	63 AAILELGIWIGNGNFIEDTEMIGAKRWIMWMMRACWPIYITILLIAIFMSLVQF 122	
DB	644 IAIPELVGISYVVGLOFCEIDIMIGFQNPNI---WKVCMAEVPITLITLFCFSFYQM 700	





```

1  APPLICANT:  Hillan, Kenneth J.
2  APPLICANT:  Phillips, Heidi
3  APPLICANT:  Polakits, Paul
4  APPLICANT:  Spencer, Susan
5  APPLICANT:  Williams, P. Mickey
6  APPLICANT:  Wu, Thomas
7  APPLICANT:  Zhang, Zhenlin
8  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
9  TITLE OF INVENTION:  TREATMENT OF TUMOR
10 FILE REFERENCE:  P5010R1-US
11 CURRENT APPLICATION NUMBER:  US/10/241,220
12 CURRENT FILING DATE:  2002-12-13
13 NUMBER OF SEQ ID NOS:  120
14 SEQ ID NO 109
15 LENGTH:  620
16 TYPE:  PRT
17 ORGANISM:  Homo Sapien
18 US-10-241-220-109

```

Query Match	26.6%	Score 319;	DB 12;	Length 620;
Best Local Similarity	36.2%	Pred. No. 6.2e-24;		
Matches 63;	Conservative 32;	Mismatches 75;	Indels 4;	Gaps 2

3 STITVTTTQDLFPKVMKKMRVPITLGGCLVLFFLLGLVCVTQAGIYVWHLIDHFCACWGILI 62  
DB 428 ESVITGSLIDEF-QLLHRHRELFTFLIVATPFLLSFCVTNGGIYVETLLDHFAGTSILE 486

487 GVLIEAIGVAMFYGVGQFSDDIDQGTGR--SLYMRCLKKLVSPECLLFVVVVSIVTF 543

Db

544 RPHYGAYIFPDMANALGNVIATSSMANPIYAAYKECSLPGSFRELAYAIAP 597

```

RESULT 6
US-09-845-908-11
: Sequence 11, Application US/09845908
: Publication NO. US20030143729A1
: GENERAL INFORMATION:
: APPLICANT: Smith, Kell E. et al
: TITLE OF INVENTION: DNA Encoding Taurine and GABA Transporters and Uses
: FILE REFERENCE: 40558-D
: CURRENT APPLICATION NUMBER: US/09/845,908
: CURRENT FILING DATE: 2001-04-30
: NUMBER OF SEQ ID NOS: 28
: SOFTWARE: PatentIn Ver. 2.0 - beta
: SEQ ID NO 11
: LENGTH: 599
: TYPE: PRT
: ORGANISM: Rattus norvegicus
: US-09-845-908-11

```

Query Match	25.88;	Score 308.5;	DB 12;	Length 599;
Best Local Similarity	31.28;	Pred.	No. 6.9e-23;	
Matches	67;	Consensus	44.4;	

3 ETITTTIIDLPPKVMKKMNPITLGGCVLELLGLVCTOAGIYVHLLIDHPCA-GWGLL 61

402 AGG TAAVADVEXPRLLRNRRRELEIAAVCAIVSYLGLSLNTTGGGIVFKEPDIYYSASGMSLL 461

122 FHRPNYGAIRPPDMGVALGKCMIVFCIIPIPIAIIKIIOANGNIFORLISCCRPASMWG 187  
       | : | | : | | : : : |  
 519 MEDTMCVNDIKKK 186

182 PYLEQNGRGERYKDMVVPKKEAGHEIPTVSGSRKPE 216  
:|:| | | | | |

Db 575 -----EDIVRP--ENGPEQPQAGSSASKE 596

RESULT 7  
US-10-345-680-5  
; Sequence 5, Application US/10345680  
; Publication No. US20030148394A1  
GENERAL INFORMATION

APPLICANT: Millennium Pharmaceuticals, Inc.  
 APPLICANT: Silos-Santiago, Inmaculada  
 APPLICANT: Venkateswarlu, Karthikeya

TITLE OF INVENTION:	UROLOGICAL DISORDERS USING FOR TREATING
TITLE OF INVENTION:	METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION:	641, 260, 55089, 21407, 42032, 46656, 52553
TITLE OF INVENTION:	34021, 44099, 25278

FILE REFERENCE: MP102-012P1RM.OMNI  
CURRENT APPLICATION NUMBER: US/10/345,680  
CURRENT FILING DATE: 2009-05-27  
FILE REFERENCE: MP102-012P1RM.OMNI  
CURRENT APPLICATION NUMBER: US/10/345,680  
CURRENT FILING DATE: 2009-05-27

PRIOR APPLICATION NUMBER: US 60/349,511  
PRIOR FILING DATE: 2002-01-18  
PRIOR APPLICATION NUMBER: US 60/360,500

PRIOR APPLICATION NUMBER: US 60/365,041  
PRIOR FILING DATE: 2002-03-15

PRIOR FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: US 60/403,468  
PRIOR FILING DATE: 2002-09-14

SENIOR APPLICATION NUMBER: US 60/414,262  
 PRIOR FILING DATE: 2002-09-27  
 PRIOR APPLICATION NUMBER: US 60/419,986

PRIOR APPLICATION NUMBER: US 60/423,809  
PRIOR FILING DATE: 2002-11-05  
PRIOR APPLICATION NUMBER: US 60/423,809

PRIOR FILING DATE: 2002-11-26  
NUMBER OF SEQ ID NOS: 66  
SOFTWARE: FastSeq for Windows Version 4

LENGTH: 599  
TYPE: PRT  
ORGANISM: Homo sapiens

S-10-345-680-5	24.7%.	Score 386
Query Match		

Matches	57; Conservatively	32.0%; Pred. NO. 1.3e-21;
	37; Mismatches	80; Indels
		4; Gaps
		2

402 EGFITALVDEYPRLLRNRRLEFIAVCCIISYLIGLSNITGGIYFVKLFDDYYSASGMSLL 461

462 FLVEFECSISWFGVNRFDYNDIQEMVGSRPCI--WWKLCWSEFTPIIYAGVFIFESAVQ 518

519 MTPLTMGNVVEPKMGCGVGMALASSMVLIPGYMAYMFLALKSLKRIQVWVPSSED 576

RESULT 8  
-09-795-693-24

Patent No. US20020068710A1  
Application US/09/35693  
GENERAL INFORMATION:  
ADDITONAL INFORMATION:

FILE REFERENCE: 35900/000303

CURRENT APPLICATION NUMBER: 05/09/795,693  
CURRENT FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: 60/185,906

PRIOR FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 24  
LENGTH: 624  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Pfam consensus sequence  
US-09-795-693-24

Query Match 24.7%; Score 295.5; DB 9; Length 624;  
Best Local Similarity 32.5%; Pred. No. 1.5e-21;  
Matches 62; Conservative 43; Mismatches 61; Indels 25; Gaps 6;

QY 3 ETTTITDLPKVKKKRVP1-TLGCCLVFLGLVCYQAGIYWHLDHCA-GWGI 60  
DB 442 EGIITAVDEPILLRKVRRELFILVCVIFSLGLFVWTEGGIYVFTLFEDYTAASGFSL 501  
QY 61 LIAAILELVGIIWYIGNRFTEDTEMIGAKRMIFLWNRACWFIPIILIAFIWSLV 120  
DB 502 LEVVEFEICIAVAVYGDIFDYDITEMLGRP---GLYKICWKVFSPLILFLFISIV 558  
QY 121 OFHRPNYGAIP-----YPMGVALGCMIVFCIIMPIIMAIITKIQAAG-N 165  
DB 559 Q-----YGLKPLTYNNWIKAEADYYPNNMANLGMILALSSMLCVPIIYIKLLSTBGS 613  
QY 166 IFQRLISCCRP 176  
DB 614 LIERLOKATTP 624

## RESULT 9

US-10-156-239-24  
Sequence 24, Application US/10156239  
Publication No. US20030036074A1  
GENERAL INFORMATION:  
APPLICANT: Glucksmann, Maria A.  
TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human Trans  
TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, A Hu  
TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor  
FILE REFERENCE: 35800/247645  
CURRENT APPLICATION NUMBER: US/10/156,239  
PRIOR FILING DATE: 2002-05-24  
PRIOR APPLICATION NUMBER: 09/795,693  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: 09/809,906  
PRIOR FILING DATE: 2000-02-29  
PRIOR APPLICATION NUMBER: 09/809,557  
PRIOR FILING DATE: 2001-03-15  
PRIOR APPLICATION NUMBER: 60/192,018  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: 09/808,568  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/191,790  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: 09/808,767  
PRIOR FILING DATE: 2001-03-15  
PRIOR APPLICATION NUMBER: 60/191,781  
PRIOR FILING DATE: 2000-03-24  
NUMBER OF SEQ ID NOS: 60  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 24  
LENGTH: 624  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Pfam consensus sequence  
US-10-156-239-24

Query Match 24.7%; Score 295.5; DB 15; Length 624;  
Best Local Similarity 32.5%; Pred. No. 1.5e-21;

Matches 62; Conservative 43; Mismatches 61; Indels 25; Gaps 6;

QY 3 ETTTITDLPKVKKKRVP1-TLGCCLVFLGLVCYQAGIYWHLDHCA-GWGI 60  
DB 442 EGIITAVDEPILLRKVRRELFILVCVIFSLGLFVWTEGGIYVFTLFEDYTAASGFSL 501  
QY 61 LIAAILELVGIIWYIGNRFTEDTEMIGAKRMIFLWNRACWFIPIILIAFIWSLV 120  
DB 502 LEVVEFEICIAVAVYGDIFDYDITEMLGRP---GLYKICWKVFSPLILFLFISIV 558  
QY 121 OFHRPNYGAIP-----YPMGVALGCMIVFCIIMPIIMAIITKIQAAG-N 165  
DB 559 Q-----YGLKPLTYNNWIKAEADYYPNNMANLGMILALSSMLCVPIIYIKLLSTBGS 613  
QY 166 IFQRLISCCRP 176  
DB 614 LIERLOKATTP 624

## RESULT 10

US-10-199-485-24  
Sequence 24, Application US/10199485  
Publication No. US20030077626A1  
GENERAL INFORMATION:  
APPLICANT: Glucksmann, Maria A.  
TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and  
TITLE OF INVENTION: 32613, No. US20030077626A1el Human Transporters  
FILE REFERENCE: 35800/249468  
CURRENT APPLICATION NUMBER: US/10/199,485  
PRIOR FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: 09/795,693  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: 60/185,906  
PRIOR FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 24  
LENGTH: 624  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Pfam consensus sequence  
US-10-199-485-24

Query Match 24.7%; Score 295.5; DB 15; Length 624;  
Best Local Similarity 32.5%; Pred. No. 1.5e-21;  
Matches 62; Conservative 43; Mismatches 61; Indels 25; Gaps 6;

QY 3 ETTTITDLPKVKKKRVP1-TLGCCLVFLGLVCYQAGIYWHLDHCA-GWGI 60  
DB 442 EGIITAVDEPILLRKVRRELFILVCVIFSLGLFVWTEGGIYVFTLFEDYTAASGFSL 501  
QY 61 LIAAILELVGIIWYIGNRFTEDTEMIGAKRMIFLWNRACWFIPIILIAFIWSLV 120  
DB 502 LEVVEFEICIAVAVYGDIFDYDITEMLGRP---GLYKICWKVFSPLILFLFISIV 558  
QY 121 OFHRPNYGAIP-----YPMGVALGCMIVFCIIMPIIMAIITKIQAAG-N 165  
DB 559 Q-----YGLKPLTYNNWIKAEADYYPNNMANLGMILALSSMLCVPIIYIKLLSTBGS 613  
QY 166 IFQRLISCCRP 176  
DB 614 LIERLOKATTP 624

## RESULT 11

US-09-815-923-10  
Sequence 10, Application US/09815923  
Publication No. US20020197644A1  
GENERAL INFORMATION:  
APPLICANT: Gill, Sarjeet S.  
APPLICANT: Ross, Linda S.

```

; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644A1e1
; FILE REFERENCE: 023070-093800US
; CURRENT APPLICATION NUMBER: US/09/815,923
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Manduca sexta
; FEATURE:
; OTHER INFORMATION: neurotransmitter transporter encoded by inebriated
; US-09-815-923-10

```

```

Query Match          24.0%; Score 287.5; DB 10; Length 727;
Best Local Similarity 35.0%; Pred. No. 1.1e-20;
Matches 63; Conservative 36; Mismatches 74; Indels 7; Gaps 3;

```

```

QY 3 ETTTITQDLFPKVMKRV---PITLGGCLVFLGLVCVTOAGIYVWHLIDHFCAGWG 59
D 525 EVVVTSIDGFPDMIRKRLVYHELVLCVAVSLGCPHIHSGIYFQMDYVAAASLS 584
QY 60 ILIAALELVGIWIGNRFIEDTEMIGAKRWIFWLMWRCWVITPILIAFIWLSL 119
D 585 ITLAFEEVVAIAMFVGVRSLRNKQMTGRP---SLYFRCCWIASPALLALMWASM 641
QY 120 VQFHRPNYGAIPYPMGVALGCMIVFCIIPMAIKIIQAKG-NIFORLISCRPAS 178
D 642 VQYTPSYRQYQYPAWQAQGLWIMASLSLSCIPYAVAVIIRABDSLRKLRYSIOPTS 701

```

```

RESULT 12
; US-09-845-908-4
; Sequence 4, Application US/09845908
; Publication No. US20030143729A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Kelli E. et al
; TITLE OF INVENTION: DNA Encoding Taurine and GABA Transporters and Uses
; FILE REFERENCE: 40558-D
; CURRENT APPLICATION NUMBER: US/09/845,908
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 4
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-845-908-4

```

```

Query Match          23.9%; Score 286.5; DB 12; Length 627;
Best Local Similarity 34.3%; Pred. No. 1.2e-20;
Matches 62; Conservative 39; Mismatches 73; Indels 7; Gaps 4;

```

```

QY 3 ETTTITQDLFPKVMKRV---PITLGGCLVFLGLVCVTOAGIYVWHLIDHFCAGWG 59
D 411 ESLVTAVDMYPRYFRGRYRELLILALSVSFELVMTLEGMTIFOLFDSYASGMC 470
QY 60 ILIAALELVGIWIGNRFIEDTEMIGAKRWIFWLMWRCWVITPILIAFIWLSL 119
D 471 LTVAFIECYCICMIVGNSNRFYDIEDMIGYRFLSIK---CMKVVTPGICAGIFIFL 527
QY 120 VQFHRPNYGAIPYPMGVALGCMIVFCIIPMAIKIIQAKG-NIFORLISCRPAS 178
D 528 VKKPKLYNNVYTPYPMGIGIGWMLALSSMLCIPILWIFIKLWKTBEIPKLOKLTVSPA 587
QY 179 N 179
D 588 D 588

```

```

RESULT 13
; US-09-845-908-2
; Sequence 2, Application US/09845908
; Publication No. US20030143729A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Kelli E. et al
; TITLE OF INVENTION: DNA Encoding Taurine and GABA Transporters and Uses
; FILE REFERENCE: 40558-D
; CURRENT APPLICATION NUMBER: US/09/845,908
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 2
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-845-908-2

```

```

Query Match          23.7%; Score 283.5; DB 12; Length 602;
Best Local Similarity 33.7%; Pred. No. 2.3e-20;
Matches 61; Conservative 40; Mismatches 73; Indels 7; Gaps 4;

```

```

QY 3 ETTTITQDLFPKVMKRV---PITLGGCLVFLGLVCVTOAGIYVWHLIDHFCAGWG 59
D 396 ESLVTALVDMYPRYFRGRYRELLILVSVSFELVMTLEGWVFOFLDYVAAASGC 455
QY 60 ILIAALELVGIWIGNRFIEDTEMIGAKRWIFWLMWRCWVITPILIAFIWLSL 119
D 456 LTVAFIESLCVAVMYGASRFNDIEDMIGYRFLP---IKTCWLFPPACVLAFFLSL 512
QY 120 VQFHRPNYGAIPYPMGVALGCMIVFCIIPMAIKIIQAKG-NIFORLISCRPAS 178
D 513 IKTPPLTKKYYTPYPMGVALGWLALSSMVICPAMSIYKLTGKPLRRLQLVCPAE 572
QY 179 N 179
D 573 D 573

```

```

RESULT 14
; US-09-815-923-4
; Sequence 4, Application US/09815923
; Publication No. US20020197644A1
; GENERAL INFORMATION:
; APPLICANT: Gill, Sarjeet S.
; APPLICANT: Ross, Linda S.
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644A
; FILE REFERENCE: 023070-093800US
; CURRENT APPLICATION NUMBER: US/09/815,923
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Manduca sexta
; FEATURE:
; OTHER INFORMATION: serotonin transporter
; US-09-815-923-4

```

```

Query Match          23.5%; Score 281.5; DB 10; Length 587;
Best Local Similarity 30.9%; Pred. No. 3.6e-20;
Matches 54; Conservative 34; Mismatches 82; Indels 5; Gaps 2;

```

```

QY 3 ETTTITQDLFPKVMKRV---PITLGGCLVFLGLVCVTOAGIYVWHLIDHFCAGWG 62
D 402 EAVTTALCDEYPRVLRHREVEVAVLLLFYICALPTTYGGVIVYLDLVNYPGLAIFL 461
QY 63 AAILLELVGIWIGNRFIEDTEMIG-AKRWIFWLMWRCWVITPILIAFIWLSLQV 121

```

```

Db      462 VVFAEAGVCWYGVDRSEEDVRLTMIGHIPGM-----FWRTCKSYISPVFELLVLFVFSVLA 517
QY      122 FHRPNYGAIPYDPWGVALGCMIVFCIIMPIMAIKIIOAKGNIFORLISCCRP 176
      518 HEMELGGEYTPSWSTVGVWMTGTIVSCIPLYIYIKLITPNCINRIKTIORP 572

```

RESULT 15

```

US-09-843-598-11
; Sequence 11, Application US/09843598
; Patent No. US2002001094A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Ranganathan, Rajesh
; TITLE OF INVENTION: CESERT GENES, PROTEINS, AND MODULATORY
; FILE REFERENCE: 01997/525002
; CURRENT APPLICATION NUMBER: US/09/843,598
; PRIOR FILING DATE: 2001-04-26
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-843-598-11

```

Query Match 23.4%; Score 280; DB 9; Length 622;

Best Local Similarity 31.9%; Pred. No. 5.5e-20; Matches 60; Conservative 34; Mismatches 90; Indels 4; Gaps 2;

```

QY      3 EITTTTIDLPKXMKMRVPTITGCGCLVFLGIVCVTOAGIYVHLDHFCAGWGILI 62
      435 EAMITALCDEPRVYGRRELFLVLLAFITLCAIPMTYGVVLFVNLVYGPGLALIF 494
QY      63 AAILLVGIIWYGNRIEDTMMIGAKRMIFWLMWRACFVITPILIAIFWISLVOF 122
      495 VVFEAGVFWYGVDRSSDVEQMLGSKPGLF--WRICWTYISPVFLTIFLFSIMGY 551
QY      123 HRPNTGAIPYDPWGVALGCMIVFCIIMPIMAIKIIOA-KGNIFORLISCCRPASNMG 181
      552 KEMLEERYYPDMYSTQGVANVICSVLCPMTIITIKFFFAKGGGRGLQESFQPEDNCG 611
Db      182 PYLEQHRG 189
QY      612 SVVPGGGG 619
Db

```

Search completed: August 27, 2003, 17:45:30  
Job time : 55 secs